

Gerau, Dieter

Disbui Qu Mainz

1989



Aus der Psychiatrischen Klinik
der Johannes-Gutenberg-Universität Mainz

Betreff:

Professor Dr. med. H. Thelen

1. Gutachter:

Professor Dr. med. H. Phillips

2. Gutachter:

Professor Dr. med. O. Bockert

Tag der

Untersuchungen zur Reliabilität und Validität
operationaler Diagnosen in der Psychiatrie
auf der Grundlage von Krankenblattauswertungen

Inauguraldissertation
zur Erlangung des Doktorgrades der Medizin
der Johannes-Gutenberg-Universität Mainz
dem Fachbereich Medizin vorgelegt

von Dieter Gerau
aus Hassloch/Pfalz

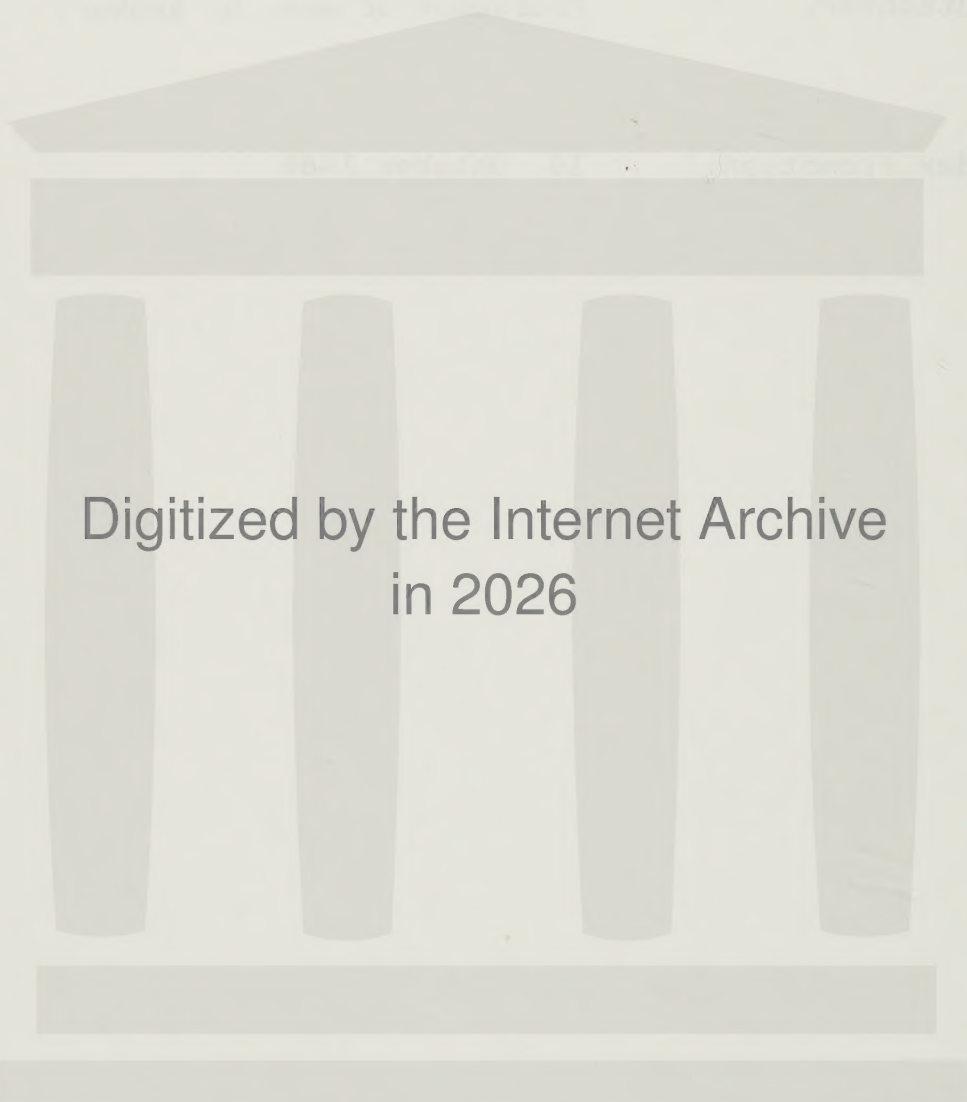
Mainz 1989

Inhaltsverzeichnis

Dekan:	Professor Dr.med. M. Thelen	Seite
1. Gutachter:	Professor Dr.med. M. Philipp	5
2. Gutachter:	Professor Dr.med. O. Benkert	5

Tag der Promotion: 18. Oktober 1989

1.	Einleitung	1
1.1.	Thema und Problemstellung	1
1.2.	Begründung des GPO-Aufbaus	2
1.3.	Ergebnisse von bisher durchgeführten vergleichenden Studien	10
1.4.	Einblick und Kritik an bisheriger Studien	12
1.5.	Grundgedanken der GPO-Erstellung	15
2.	Methodik	16
2.1.	Patientenerschöpfung	16
2.2.	Vorstellung der Klassifikationssysteme	18
2.2.1.	Research Diagnostic Criteria (RDC)	18
2.2.2.	Diagnostic and Statistical Manual of Mental Diseases (DSM-III)	19
2.2.3.	Feighner-Kriterien (FCO)	21
2.2.4.	Newcastle Scores I & II: endogene Depression	21
2.3.	Interviewdiagnostik	21
2.4.	Krankenblattdiagnostik	22
2.5.	Datenverarbeitung	24
2.5.1.	Statistische Verfahren	24
2.5.1.1.	Statistische Methodik zur Erfassung der Reliabilität	25
2.5.1.1.1.	Berechnung des Kappa-Koeffizienten	26
2.5.1.1.2.	Untersuchung der Validität	27
2.5.1.1.2.1.	Berechnung der Sensitivität	29
2.5.1.1.2.2.	Berechnung der Spezifität	29
2.5.1.1.2.3.	Der positiv prädiktive Wert (PPV) bei	30
3.	Ergebnisse	31
3.1.	Reliabilität	31
3.1.1.	Reliabilität auf Diagnoseebene	31
3.1.2.	Reliabilität auf Symptom-Ebene	35
3.1.2.1.	Depressive Erkrankung	35
3.1.2.2.	Unipolares depressives Syndrom	36
3.1.2.3.	Bipolare Erkrankung	36
3.2.	Ubereinstimmungsvalidität	38



Digitized by the Internet Archive
in 2026

https://archive.org/details/IA41554422_0052

Inhaltsverzeichnis

	Seite
1. <u>Einleitung</u>	5
1.1. Technik und Aufbau der operationalisierten Diagnostik (OPD)	6
1.2. Beispiel des OPD-Aufbaus	8
1.3. Ergebnisse von bisher durchgeführten vergleichenden Studien	10
1.4. Methodik und Ergebnisse bisheriger Studien auf der Grundlage von Krankenblattauswertungen	12
1.5. Fragestellung	15
2. <u>Methodik</u>	16
2.1. Patientenstichprobe	16
2.2. Vorstellung der Klassifikationssysteme	18
2.2.1. Research Diagnostic Criteria (RDC)	18
2.2.2. Diagnostic and Statistical Manual of Mental Diseases (DSM-III)	19
2.2.3. Feighner-Kriterien (FDC)	21
2.2.4. Newcastle Scales I & II: endogene Depression	21
2.3. Interviewdiagnostik	22
2.4. Krankenblattdiagnostik	22
2.5. Datenverarbeitung	24
2.5.1. Statistische Verfahren	24
2.5.1.1. Statistische Methodik zur Erfassung der Reliabilität	25
2.5.1.1.2. Berechnung des Kappa-Koeffizienten	26
2.5.1.2. Untersuchung der Validität	27
2.5.1.2.1. Berechnung der Sensitivität	29
2.5.1.2.2. Berechnung der Spezifität	29
2.5.1.2.3. Der positiv prädiktive Wert (ppW 50)	30
3. <u>Ergebnisse</u>	31
3.1. Reliabilität	31
3.1.1. Reliabilität auf Diagnosenebene	31
3.1.2. Reliabilität auf Symptom-Ebene	35
3.1.2.1. Depressive Erkrankung	35
3.1.2.2. Übrige depressive Syndrome	40
3.1.2.3. Manische Erkrankung	44
3.2. Übereinstimmungsvalidität	46

	Seite
4. <u>Diskussion</u>	53
4.1. Diskussion der Ergebnisse der Reliabilitätsprüfung	53
4.1.1. Die diagnostischen Hauptkategorien	54
4.1.2. Die diagnostischen Untergruppen	57
4.1.2.1. Die endogene Depression	57
4.1.2.2. Übrige depressive Syndrome	59
4.2. Diskussion der Ergebnisse der Validitätsprüfung	60
4.2.1. Die diagnostischen Hauptkategorien	60
4.2.2. Die diagnostischen Untergruppen	62
5. <u>Zusammenfassung</u>	64
6. <u>Tabellarischer Anhang</u>	67
7. <u>Literaturverzeichnis</u>	81

1. Einleitung

Es wird weithin akzeptiert, daß in der psychiatrischen Forschung standardisierte und reliable Diagnosen benötigt werden, um Diagnosenträger reproduzierbar zu identifizieren.

Insbesondere die biologische Forschungsrichtung der Psychiatrie ist auf diagnostische Verfahrensweisen angewiesen, die in der Lage sind, homogene Patientengruppen mit hoher diagnostischer Reliabilität zu charakterisieren.

In den letzten Jahren hat es daher zahlreiche Bemühungen gegeben, psychiatrische Diagnosen zu operationalisieren und auf diesem Weg dem gesteckten Ziel etwas näherzukommen. Hauptvorteil solcher operationalen Diagnosen (OPD) ist ein Anstieg der diagnostischen Reliabilität und eine bessere Vergleichbarkeit der Patientengruppen in klinischen Studien (Spitzer et al. 1978, Kendell 1982).

Die bekanntesten OPD sind: die Feighner-Kriterien (Feighner, Robins und Guze 1972), die Research Diagnostic Criteria (RDC; Spitzer et al. 1975 und 1978), die Taylor-Abrams-Kriterien (Abrams et al. 1974, Taylor et al. 1974 und 1978), die Wiener Forschungskriterien (Berner et al. 1983) und das Diagnostic and Statistical Manual of Mental Disorders, Third Edition (DSM III) der American Psychiatric Association (APA 1980); Revision DSM-III-R (APA 1987).

Diese Systeme werden nicht nur für die Forschung, sondern zunehmend auch im Klinikalltag verwendet; DSM III und DSM-III-R beispielsweise sind ausdrücklich als ein versorgungsbezogenes Diagnosesystem konzipiert.

1.1. Technik und Aufbau der operationalisierten Diagnostik

Sinn der operationalisierten Diagnostik ist es, eine einheitliche und eindeutige Definition der diagnostischen Kriterien zu gewährleisten und dadurch die diagnostische Übereinstimmung (Reliabilität) zwischen verschiedenen Untersuchern zu erhöhen.

Daß die Reliabilität tatsächlich durch Anwendung operationaler Diagnosekriterien erhöht werden kann, wurde für die genannten Diagnoseschemen mehrfach gezeigt (Spitzer et al. 1978; Rounsaville et al. 1980; Klerman, Endicott & Spitzer 1979; Wittchen 1985; Fuhrer 1986).

Diese Befunde beziehen sich allerdings ausnahmslos auf eine Diagnostik mittels standardisierten Interviewverfahren wie Present-State-Examination (PSE; Wing et al. 1974), Diagnostic Interview Schedule (DIS; Robins et al. 1981), Schedule for Affective Disorders and Schizophrenia (SADS; Endicott und Spitzer 1978) oder Structured Clinical Interview for DSM III (SCID; Spitzer 1983; Spitzer und Williams 1983).

Hierdurch soll systematisch das Vorliegen bzw. das Fehlen der einzelnen Kriterien festgestellt werden.

Die strukturierte Befragung des Patienten ist somit bei diesem Vorgehen die wichtigste Informationsquelle.

Es gibt allerdings auf der anderen Seite einige Publikationen in relevanten Journalen, in welchen über eine operationalisierte Diagnosestellung aufgrund der Auswertung von Krankenblatttexten berichtet wird (Zaudig und Vogel 1983, Friedman et al. 1982, Craig et al. 1982, Silverstein et al. 1982, Helzer et al. 1978, Nelson et al. 1981, Winokur 1985, Marneros et al. 1988 a-c)



Als Vorteile der Krankenblattauswertung gegenüber dem live-Rating werden u.a. angeführt die bessere Auswahlmöglichkeit und damit verbunden eine repräsentative Verteilung der Diagnosen, die leichte Verfügbarkeit großer Fallzahlen und die Tatsache, daß ohne wesentlich größeren Aufwand verschiedene miteinander konkurrierende Diagnosesysteme angewandt werden können. Eine ausschließliche Krankenblattauswertung geht aber gleichzeitig von der impliziten Annahme aus, daß die so erhaltenen Ergebnisse denen von live-Ratings vergleichbar sind, also somit die gleiche Reliabilität und Validität erreicht wird. Ob diese Vergleichbarkeit denn überhaupt gegeben sei, hatten bereits Hylert et al. 1982 in Frage gestellt.

Bislang ist allerdings noch nicht untersucht worden, ob eine operationalisierte Diagnostik alleine durch eine Krankenblattauswertung mit ausreichender Reliabilität und Validität erhoben werden kann.

Es müssen sicherlich Zweifel an der Erreichbarkeit eines solchen Zieles geäußert werden, wenn man an die Komplexität und Differenziertheit einzelner Kriterien denkt, wie z.B. Früherwachen mindestens eine Stunde vor der gewohnten Zeit, Gewichtsverlust von mindestens 0,5 kg pro Woche, besondere Qualität der depressiven Verstimmung, tageszeitliche Schwankungen der depressiven Symptomatik, unangemessene Schuldgefühle u.a.m.

1.2. Beispiel des OPD-Aufbaus

Um diese soeben angesprochene Komplexität der Kriterienlisten an dieser Stelle etwas anschaulicher vor Augen zu führen, sollen im folgenden zunächst repräsentativ für den Aufbau aller operationalisierten Diagnosesysteme (OPD) die Kriterien des DSM-III, "Major Depressive Episode (MDE) with melancholia", denen der Research Diagnostic Criteria (RDC), "Major Depressive Disorder (MDD), endogenous type" gegenübergestellt werden.

Tabelle 1) RDC, Major Depressive Disorder, endogenous type

I und II müssen erfüllt sein:

- I) Die Kriterien einer Major Depressive Disorder (MDD) sind erfüllt
 - II) Mindestens ein Kriterium der Gruppe A) und mindestens sechs aus A) und B) sind erfüllt
-

- A1) Besondere Qualität der depressiven Verstimmung, d.h. die depressive Gestimmtheit wird als etwas eindeutig anderes erlebt als jene Stimmung, die nach dem Tode einer geliebten Bezugsperson erlebt wird
 - A2) Fehlende Reaktivität auf Veränderung der Umgebung d.h. während der Depression fühlt sich der Patient auch dann nicht einmal zeitweise besser, wenn sich etwas Erfreuliches ereignet
 - A3) Die Stimmung ist regelmäßig morgens schlechter
 - A4) Durchgängiger Verlust von Interesse oder Vergnügen
 - B1) Selbstvorwürfe oder ausgeprägte oder inadäquate Schuldgefühle
 - B2) Frühes morgendliches Erwachen oder Durchschlafstörungen
 - B3) Objektivierbare psychomotorische Verlangsamung oder Agitation
 - B4) Appetitmangel
 - B5) Gewichtsverlust ($\geq 1\text{kg/Woche}$ über mehrere Wochen oder $\geq 10\text{kg/1 Jahr}$)
 - B6) Interesseverlust oder Verlust an Vergnügen im Bereich üblicher Aktivitäten oder vermindertes sexuelles Interesse
-

Tabelle 2) DSM-III, MDE with melancholia

I und II müssen erfüllt sein

I) Die Kriterien einer MDE sind erfüllt

II) A1 und A2, sowie mindestens drei Kriterien von B1 bis B6 treffen zu

A1) Verlust des Vergnügens im Bereich aller oder fast aller Aktivitäten

A2) Fehlende Reaktivität auf normalerweise vergnügliche Ereignisse, fühlt sich bei vergnüglichen Ereignissen nicht einmal vorübergehend besser

B1) Besondere Qualität der depressiven Verstimmung, d.h. die depressive Gestimmtheit wird als etwas eindeutig anderes erlebt als jene Stimmung, die nach dem Tod einer geliebten Bezugsperson erlebt wird

B2) Die Stimmung ist morgens regelmäßig schlechter

B3) Früherwachen von mind. 2 Stunden vor der gewohnten Aufwachzeit

B4) Ausgeprägte psychomotorische Hemmung oder Agitation

B5) Gravierende Appetitlosigkeit oder Gewichtsminderung

B6) Exzessive oder unangemessene Schuldgefühle

Beiden operationalisierten Diagnoseschemen ist zunächst gemeinsam, daß ihre jeweiligen Vorbedingungen – also das sichere Vorliegen einer Major Depressive Disorder bzw. einer Major Depressive Episode – erfüllt sind.

Sie unterscheiden sich jedoch in der Restriktivität der Bedingungen, die zu erfüllen sind, wenn von endogener Depression die Rede sein soll. Die Kriterienliste der Research Diagnostic Criteria (RDC) erscheint dabei weniger restriktiv als die des Diagnostic and Statistical Manual of Mental Diseases (DSM-III).

Im Falle der RDC-Kriterien werden zwar letztlich insgesamt sieben verschiedene Items zur Diagnosestellung gefordert, während im DSM-III nur vier notwendig sind. Allerdings bietet RDC die Option, aus insgesamt zehn Items auswählen zu können, wobei nur ein einziges Kriterium der Gruppe A), die die restriktiven Kriterien enthält, ausreichen würde, um die Diagnose zu stellen.

Im DSM-III werden insgesamt nur acht Items angeboten, von denen beide der Gruppe A) erfüllt sein müssen. Damit steht das Anhedonie-Konzept als obligatorisches Symptom im Zentrum dieser Klassifikation (Maier, Philipp und Buller 1986).

In den bisherigen Publikationen, die sich mit dem polydiagnostischen Ansatz im Vergleich verschiedener OPD befassen, waren sich die Autoren übereinstimmend darin einig, daß im Falle der operationalisierten Diagnosekriterien des DSM-III eine deutlich ausgeprägtere Restriktivität herrscht, als z.B. in der Operationalisierung der RDC (Maier et al. 1986, Copolov et al. 1986, Philipp und Maier 1985, Davidson et al. 1984a).

1.3. Ergebnisse von bisher durchgeführten vergleichenden Studien

In dem Bemühen die Reliabilität psychiatrischer Diagnosen zu erhöhen, die zuvor laut Spitzer und Fleiss (1974) nur unbefriedigend niedrige Werte aufwies, begann man seit Anfang der siebziger Jahre mit dem Aufbau operationalisierter Diagnoseverfahren (OPD), für die zum Teil sogar Computer-gestützte Verfahren entwickelt wurden wie z.B. PSE-Catego (Bronisch et al. 1982). Bis zum gegenwärtigen Zeitpunkt dürften in etwa ein Dutzend miteinander konkurrierende Modelle publiziert worden sein, deren Fähigkeit, die Reliabilität psychiatrischer Diagnosen zu erhöhen, erwiesen ist (vgl. 1.1 und 1.2.).

Gleichzeitig mit der Verbreitung dieser OPD wurden auch Untersuchungen durchgeführt, um zum einen die jeweilige Validität zu überprüfen, zum anderen aber auch um die verschiedenen OPD auf ihre Vergleichbarkeit hin zu untersuchen. In bezug auf die Validität wurden als Validierungskriterium teilweise biologische Marker, wie der Behandlungserfolg nach Antidepressiva-Gabe (z.B. Stewart et al. 1983) und der Dexamethason-Suppressionstest (z.B. Zimmerman et al. 1985b, Davidson et al. 1984b), teilweise die Übereinstimmung mit den klinischen ICD-9-Diagnosen (Philipp et al. 1985a; Maier et al. 1986c) benutzt.

Bei den Vergleichsuntersuchungen der einzelnen OPD miteinander ließen sich deutliche Unterschiede feststellen.

So fanden Copolov et al. (1986) in der vergleichenden Klassifikation von 70 depressiven Patienten nach RDC und DSM-III deutliche Unterschiede, die auf die höhere Restriktivität der DSM-Kriterien zurückgeführt wurde. Alle Patienten, die darin nach DSM-III die Kriterien der Major Depressive Episode (MDE) with melancholia erfüllten, erfüllten auch die Bedingungen der Research Diagnostic Criteria (RDC) endogenous type. Umgekehrt waren aber nur 20 von 35 Patienten, die nach RDC als endogen depressiv klassifiziert worden waren, auch in DSM-III in der äquivalenten Sparte "with melancholia" klassifizierbar.

Dabei zeigten insbesondere die in Tabelle 2) unter Gruppe B) angeführten Kriterien die gleiche Häufigkeit in den Untergruppen melancholisch(DSM)/definitiv endogen(RDC) als auch in der, der als nicht melancholisch (DSM), aber sicher endogen (RDC) eingestuften Patienten. Die Autoren zogen daraus die Schlußfolgerung, daß die höhere Restriktivität der DSM-III Kriterien ausschließlich in den obligatorisch geforderten Symptomen der Gruppe A) begründet sei.

Davidson et al. (1984a) berichteten ähnliche Resultate; in ihrer vergleichenden Diagnostik-Studie von 50 depressiven Patienten konnten 50% als definitiv endogen nach

RDC, jedoch nur 30% als "melancholisch" nach DSM-III eingeordnet werden. Diese unterschiedliche diagnostische Zuordnung in den beiden Systemen wurde auch von Zimmerman et al. (1985a) bestätigt. Dort waren von 159 Patienten mit einer depressiven Erkrankung 67% nach RDC als endogen eingeordnet worden, dagegen nur 44% in DSM-III.

Noch deutlichere Unterschiede zwischen den einzelnen OPD waren in einer Studie von Philipp und Maier (1987) berichtet worden. In einer Vergleichsuntersuchung acht verschiedener OPD wurden zwar in den meisten Fällen eine ausreichend gute Assoziation für die jeweilige Kategorie "endogene Depression" gefunden, die zufallskorrigierte diagnostische Übereinstimmung war jedoch eher unbefriedigend mit Kappa-Werten kleiner als .50 . Damit liegt dieses Ergebnis nicht höher, als in der Untersuchung von Spitzer und Fleiss 1974, vor Verbreitung der OPD.

1.4. Methodik und Ergebnisse bisheriger Studien auf der Grundlage von Krankenblattauswertungen

Eine Untersuchung auf alleiniger Basis der Auswertung von Krankengeschichten war bisher eher die Ausnahme. So beziehen sich alle unter 1.3. zitierten Studien auf die Ergebnisse eines Ratings direkt am stationären Patienten.

Helzer et al. (1978) entwarfen ein Studien-Setting, in dem die diagnostischen Ergebnisse eines life-Ratings von 101 Patienten zu einem späteren Zeitpunkt mit den Diagnosen der Krankenakten dieser Patienten verglichen wurden.

Dabei wurden alle in der Krankenakte festgehaltenen Informationen verwendet. In der ursprünglichen Reliabilitätsstudie (Helzer et al. 1977b) wurden zwei voneinander unabhängig durchgeführte Ratings zur Frage der Test-Retest-Reliabilität miteinander verglichen. Nun wurden die damaligen Ergebnisse mit den jetzt erhaltenen verglichen, nachdem die aus den Krankengeschichten der gleichen Patienten gewonnenen Informationen ebenfalls operationalisiert worden waren. Der daraufhin berechnete Kappa-Koeffizient zeigte generell niedrigere Übereinstimmung zwischen diagnostischem Interview und Krankenblattauswertung als zwischen den beiden unabhängig voneinander durchgeführten Interviews.

Nelson et al. (1981) konzentrierten sich in ihrer Studie in erster Linie auf die DSM-III-Kategorie "Melancholie" (endogene Depression). Dabei wurden retrospektiv die Krankengeschichten von 127 unipolar depressiv erkrankten Patienten ausgewertet, und unter Anwendung der Schedule for Affective Disorders and Schizophrenia (SADS; Endicott & Spitzer 1978) wurden die An-/Abwesenheit der Symptome für die DSM-III Diagnosen Major Depressive Episode (MDE), MDE with melancholia und die Feighner-Kriterien für Depression beurteilt.

Es zeigte sich, daß die Melancholie-Kriterien nach DSM-III 77% der Patientenstichprobe korrekt klassifizierten und sich gleichzeitig gegenüber den Feighner-Kriterien restriktiver zeigten, die ihrerseits lediglich 67% korrekt zuordneten, was mit deren niedriger Spezifität erklärt worden war.

Besonders die Items "fehlende Reaktivität" und "veränderte Psychomotorik" des DSM-III erwiesen sich als besonders geeignet, zwischen den Syndromen "autonom" und "reaktiv" zu diskriminieren (vgl. hierzu Abschnitt 4.1.2.1.)

Eine weitere retrospektiv angelgte Studie (Craig et al. 1982) verglich die klinischen DSM-II Diagnosen von 102 zufällig ausgewählten Krankengeschichten der Jahre 1975/76 mit den Diagnosen, die sich nach den Kriterien von DSM-III ergeben würden.

Allerdings waren nur Aufnahme- und Entlassungsbefunde der Krankenblätter benutzt worden, nicht dagegen die Eintragungen zum Verlauf. Diagnosen wurden aus den Exzerpten vor Auswertung getilgt, das dokumentierte Fallmaterial wurde dann nach den Kriterienlisten des DSM-III ausgewertet. Neben dem wohl USA-spezifischen Befund, daß die Abweichungen zwischen klinischer und DSM-Diagnose für Farbige wesentlich deutlicher ausgeprägt waren als für Weiße, erscheint bemerkenswert, daß bei insgesamt 21 Patienten ein Diagnosenwandel unter DSM-III-Kriterien eintrat, und zwar in erster Linie in den Kategorien "Schizophrenie" und "affektive Störungen".

Die kappa-Berechnung zeigte jedoch statistisch signifikante Übereinstimmungen in allen Untergruppen zwischen $k=.79$ und $k=.54$.

Bestehende Abweichungen werden darauf zurückgeführt, daß nicht alle vorhandenen Symptome im Krankenblatttext festgehalten worden waren, da zum damaligen Zeitpunkt kein strukturiertes Aufnahmeinterview benutzt wurde. Die Häufigkeit der Neu-Klassifikation im Bereich der schizophrenen Erkrankungen wurde vor allem mit dem DSM-III-Kriterium "Dauer der Störung länger als 6 Monate" erklärt.

In der Untersuchung von Hyler, Williams und Spitzer (1982) war die Reliabilität der DSM-III Diagnosen beurteilt worden. Hierbei wurde die diagnostische Inter-Rater-Reliabilität von zwei "live"-Interviews (Joint- und Test-retest-Ratings) einerseits mit der zwischen Interview und Krankenblattauswertung andererseits verglichen. Das gefundene Resultat zeigte eine deutlich höhere Übereinstimmung der Diagnosen zwischen den Beurteilungen der beiden Interview-Rater als zwischen den Ergebnissen der Kombination Krankenblattauswerter vs. Interview-Rater.

1.5. Fragestellung

Die vorliegende Arbeit hat sich zur Aufgabe gestellt, Untersuchungen zur Reliabilität und Validität der operationalisierten Diagnostik aufgrund von Krankenblattaufzeichnungen durchzuführen.

Im einzelnen sollen die folgenden beiden Fragen untersucht werden:

- a. Können Krankenblattinformationen mit ausreichender Reliabilität in operationalisierte Diagnosen überführt werden?
- b. Ist auf der Basis alleiniger Krankenblattauswertungen und Überführung der darin enthaltenen Information in operationalisierte Diagnosen eine ausreichend valide Diagnostik möglich?

2. Methodik

2.1. Patientenstichprobe

Ausgangsstichprobe war eine Gruppe von n=206 Patienten, die bei stationärer Aufnahme in der Psychiatrischen Universitätsklinik Mainz zwischen Oktober 1982 und Juni 1983 mit der Present State Examination (PSE; Wing et al. 1974) von insgesamt vier erfahrenen Psychiatern standardisiert interviewt wurden.

Es handelte sich dabei um konsekutiv stationär aufgenommene Patienten, die folgenden Ein- und Ausschlusskriterien genügten:

A) Einschlusskriterien:

1. Lebensalter zwischen 20 und 60 Jahren
2. Die Patienten erklärten sich nach entsprechender Aufklärung mit der Durchführung des Interviews einverstanden, sodaß die von ihnen gewonnen psychopathologischen Daten in anonymisierter Form in verschiedenen Studien verwendet werden konnten.

B) Ausschlusskriterien:

1. Körperlich begründbare psychische Störungen mußten ausgeschlossen werden können
2. Patienten mit schweren körperlichen Erkrankungen wurden nicht in die Studie einbezogen
3. Das Vorliegen von Alkohol- und/oder Drogenabhängigkeit führte gleichfalls zum Ausschluß

Dieses Patientenkollektiv stellte das Untersuchungsgut dar, über das verschiedene psychopathometrische Verlaufsstudien bereits veröffentlicht sind (Maier et al. 1986a-f und 1987; Philipp et al. 1985a+b, 1986a-d; Philipp und Maier 1987). Dazu waren die im PSE und einem sich anschließenden freien Interviewteil gewonnenen Informationen anschließend bezüglich der Kriterienlisten folgender operationalisierter Diagnosesysteme beurteilt worden:

1. Research Diagnostic Criteria
(RDC; Spitzer et al. 1978)
2. Diagnostic and Statistical Manual of Mental Disorders (DSM-III; APA 1980)
3. Feighner-Kriterien (Feighner et al. 1972)
4. Newcastle Scales I und II
(Carney et al. 1965 und Gurney 1971)
5. Taylor-Abrams Kriterien
(Abrams et al. 1974; Taylor et al. 1974 und 1978)
6. Wiener Forschungskriterien (VRC)
(Berner et al. 1983)

Aus dieser Ausgangsstichprobe von ursprünglich n=206 Patienten wurde eine Zufallsstichprobe von n=97 Patienten gezogen. Auf diese Teilstichprobe bezieht sich die im folgenden vorgestellte Untersuchung.

Die unabhängig vom behandelnden Arzt gestellten Entlassungsdiagnosen der Patienten dieser Teilstichprobe verteilt sich nach der International Classification of Diseases, 9.Auflage (ICD-9; Diagnoseschlüssel und Glossar psychiatrischer Krankheiten, Springer Verlag 1980) wie folgt:

Tabelle 3) Charakterisierung der Ausgangsstichprobe

ICD 295.x	schizophrene Psychosen (ohne ICD 295.7)	n=26
ICD 296.x	affektive Psychosen	n=28
ICD 295.7	schizo-affektive Psychosen	n=12
ICD 300.4	neurotische Depression	n=14
ICD 300.x	andere neurot.Störungen	n=10
ICD 309.0	kurze depressive Reaktion	n= 7
Summe		n=97

In die Krankenblattauswertung eingegangen sind neben der aktuellen Symptomatik auch alle Daten, die sich auf Längsschnittsymptome beziehen wie z.B. Vorgeschichte, Bipolarität und Familienanamnese.

2.2. Vorstellung der Klassifikationssysteme

Wir entschieden uns für einen polydiagnostischen Klassifikationsansatz, wie er u.a. von Berner und Katschnig (1983), Katschnig und Egger-Zeidner (1985), Zimmerman et al. (1984, 1985, 1986) und anderen empfohlen wurde. Diese simultane Klassifikation nach miteinander konkurrierenden Diagnosesystemen erlaubt eine vergleichende Betrachtung der jeweiligen inhaltlichen Validität.

Zum Einsatz kamen vier weitverbreitete Klassifikationsschemata, die im folgenden kurz vorgestellt und charakterisiert werden sollen.

2.2.1. Research Diagnostic Criteria (RDC; Spitzer et al. 1978)

In der vorliegenden Arbeit wurden die nachfolgend tabellarisch aufgelisteten Diagnosen angewandt.

Tabelle 4) Untersuchte RDC-Diagnosen

Depressive Syndrome			Manische Syndrome	
RDC 9	Major Depressive Disorder (MDD)		RDC 5	Manische Erkrankung
RDC 9a	primäre MDD			
RDC 9b	sekundäre MDD		RDC 6	Hypomanische Erkrankung
RDC 9c	wiedekkehrende unipolare MDD			
RDC 9d	psychotische MDD			
RDC 9e	inaktivierende MDD			
RDC 9f	endogene MDD			
RDC 9g	agitierte MDD			
RDC 9h	gehemmte MDD			
RDC 9i	Situative MDD			
RDC 9j	einfache MDD			
RDC 10	Minor Depressive Disorder			
RDC 11	intermittierende depressive Erkrankung			
RDC 7	bipolare Depression mit Manie (Bipolar 1)			
RDC 8	bipolare Depression mit Hypomanie (Bipolar 2)			

Die Kriterienlisten der RDC sind hierarchisch strukturiert, es werden sowohl Ein- als auch Ausschlusskriterien definiert. Ausdrückliches Ziel ist die Bildung von diagnostisch möglichst homogenen Patientengruppen.

Jede Diagnose kann als "sicher", "wahrscheinlich" oder als "sicher nicht" vorhanden beurteilt werden. In der vorliegenden Untersuchung wurden nur die als "sicher" beurteilten Diagnosen und Symptome als vorhanden gewertet. Die klinische Diagnose einer endogenen Depression setzt das Vorliegen einer Major Depressive Disorder (MDD) voraus. Darüberhinaus müssen bestimmte zusätzliche Kriterien erfüllt sein.

Neben einer dichotomen Einteilung in eine primäre bzw. sekundäre MDD werden weitere Klassifikationsmöglichkeiten durch eine psychotische und situative MDD und – je nach vorherrschender Psychomotorik – durch eine agitierte bzw. gehemmte MDD angeboten.

Die ebenfalls beschriebene Minor Depressive Disorder findet in der gebräuchlichen ICD-9 Klassifikation keine Entsprechung. Sie unterscheidet sich von der Major Depressive Disorder durch die geringere Intensität der depressiven Verstimmung. Dementsprechend werden weniger Einzel-Symptome für diese Diagnose gefordert.

Die Intermittierende Depressive Erkrankung nach RDC entspricht am ehesten der ICD-9 Diagnose einer Neurotischen Depression (ICD 300.0).

2.2.2. Diagnostic and Statistical Manual of Mental Diseases (DSM III), Third Edition (American Psychiatric Association 1980)

Zunächst soll auch an dieser Stelle – wie schon bei den RDC Kriterien eine tabellarische Auflistung der verwendeten DSM-III-Diagnosen erfolgen.

Tabelle 5 Untersuchte DSM-III-Diagnosen

Depressive Syndrome

Manische Syndrome

Major Depressive Episode (MDE)

MDE, einzelne Episode

MDE, wiederholte Episode

Bipolare affektive Erkrankung,
depressiv

MDE with melancholia

Dysthymie

Anpassungsstörung mit depressiven Zügen

atypische Depression

Manische Episode (ME)

ME, einzelne Episode

ME, wiederholte Episode

Bipolare affektive

Erkrankung, manisch

Bipolare affektive

Erkrankung, gemischt

Cyclothyme Erkrankung

atypische bipolare

Erkrankung

Im Gegensatz zu allen anderen operationalen Diagnosesystemen verfolgt DSM-III einen multiaxialen Klassifikationsansatz. In dieser Arbeit beschränkt sich die Untersuchung jedoch auf Achse I. Wie bei RDC werden auch hier Ein- und Ausschlusskriterien definiert und durch zusätzliche Zeit- und Schweregradkriterien ergänzt. Anders als bei den übrigen Systemen sind nur definitive Diagnosen möglich.

Analog zur Major Depressive Disorder bei den RDC wird die Major Depressive Episode (MDE) definiert. Vergleichbar dem RDC wird ein endogener Untertyp der MDE angeboten, der als MDE with melancholia bezeichnet wird.

Dieses Melancholie-Konzept trägt im wesentlichen den Untersuchungen von Nelson und Charney (1980, 1981) Rechnung, die wesentlichen Nachdruck auf das Kriterium der Anhedonie legen. Daneben werden die diagnostischen Kategorien einer neurotischen Depression (Dysthymie), einer Anpassungsstörung mit depressiven Zügen und für nicht klassifizierbare Patienten die Restkategorie der atypischen Depression operationalisiert.

2.2.3. Feighner-Kriterien (FDC; Feighner et al. 1972)

Die 1972 veröffentlichten Feighner-Kriterien - auch als St. Louis-Kriterien bekannt geworden - sind als die unmittelbaren Vorläufer der späteren Research Diagnostic Criteria (RDC) anzusehen. Dieses System enthält lediglich Einschluss-Kriterien.

Die Diagnose der primären Depression wird - anders als MDD und MDE der RDC bzw. DSM-III Kriterien - nicht weiter in Untertypen aufgeteilt. Es entfällt also insbesondere eine gesondert definierte Kategorie der endogenen Depression.

2.2.4. Newcastle Scale I: endogene Depression (NC1)

(Carney et al. 1965, modifiziert durch Bech et al. 1980)

Newcastle Scale II: endogene Depression (NC2)

(Gurney 1971, modifiziert durch Bech et al. 1980)

Wurden RDC, DSM-III und FDC durch Konsens-Bildung klinischer Experten entwickelt, so lag der Konstruktion beider Newcastle-Skalen ein statistisches Verfahren, nämlich das einer Diskriminanzanalyse zugrunde. Es handelt sich also um ein grundsätzlich verschiedenes Konstruktionsprinzip. Aus einer umfangreichen Liste potentiell diskriminierender Kriterien wurden die jeweils zehn in der Diskriminanzanalyse am besten zwischen endogener und neurotischer Depression unterscheidenden Items ausgewählt und gewichtet.

Die Gewichtung dieser jeweils 10 Items werden zu einem Score aufsummiert und ein sogenannter "cut-off-Punkt" ermittelt, der am besten zwischen klinisch beurteilten endogenen und nicht-endogenen Depressionen unterscheidet. Eine weitergehende Operationalisierung durch zusätzliche Ein- bzw. Ausschlusskriterien ist nicht erfolgt.

Die von P. Bech et al. vorgenommene Modifikation der ursprünglichen Newcastle Skalen zeichnet sich aus durch eine ausführlichere Beschreibung der einzelnen Items und durch die Möglichkeit zwischen definitivem und wahrscheinlichem Zutreffen eines Items zu wählen.

2.3. Interviewdiagnostik

Jeder Patient wurde innerhalb der ersten Woche seines stationären Aufenthaltes, möglichst aber innerhalb von drei Tagen nach dem Aufnahmetermin standardisiert mit dem PSE (Present State Examination; Wing et al. 1974) interviewt.

Diese Fremdbeurteilungen wurden von einer Gruppe von vier Psychiatern mit zwischen 2 und 13 Jahren klinischer Erfahrung vorgenommen. Sämtliche Rater wurden vor Untersuchungsbeginn mit den Anforderungen eines PSE-Interviews vertraut gemacht und nahmen in der Vorphase der Untersuchung an einem Rater-Training teil. Jeder Rater hatte im Rahmen dieses Trainings an mindestens 15 Interviews teilgenommen.

In unmittelbarem Anschluss an das strukturierte PSE-Interview wurden die operationalen Diagnosesysteme RDC, DSM-III, FDC, NC1 und NC2 unter Verwendung zusätzlicher freier Fragen beurteilt.

2.4. Krankenblattdiagnostik

Der Aufbau der Krankenblatttexte ist in der Psychiatrischen Universitätsklinik Mainz normiert und somit weitgehend vergleichbar. Der Textgehalt der Krankengeschichten erstreckte sich auf zwischen 8 und 21 Maschinentextseiten mit je 400 Worten pro Seite im Mittel.

Der Textgehalt wurde als Anzahl geschriebener Worte im Belegbogen dokumentiert.

Die Auswertung der Krankenblätter erfolgte für jeden Patienten ca. 18 Monate nach den PSE-Interviews durch zwei voneinander unabhängige Untersucher, die weder mit dem Verfasser der jeweiligen Krankengeschichte noch mit dem Rater der Interviewdiagnostik (siehe Abschnitt 2.3.) identisch waren.

Krankenblattauswerter waren der Referent, sowie fünf in Ausbildung zum Arzt für Psychiatrie befindliche Assistenzärzte der Psychiatrischen Universitätsklinik Mainz. Alle waren blind bezüglich der Ergebnisse des PSE-Interviews und der daraus gewonnenen operationalisierten Diagnosen. In den Krankenblattaufzeichnungen nicht explizit erwähnte, obwohl wahrscheinlich vorhandene Symptome wurden konsequent als abwesend gewertet.

Die zur Anwendung gekommenen klassifikatorischen Kriterien wurden auf einer sechstufigen Skala bezüglich der Sicherheit des Vorliegens bzw. Nicht-Vorliegens beurteilt.

(Vgl. tabellarischer Anhang)

Die Scores 1-3 meinen hierbei die Abwesenheit des entsprechenden Kriteriums, die Scores 4-6 dessen Abwesenheit aufgrund des im Krankenblatt niedergelegten Befundes. 1 bzw. 6 sind die jeweiligen Pole der höchsten Beurteilungssicherheit, 3 und 4 die der niedrigsten.

Die Beurteilungen der beiden Rater wurden in einem weiteren Schritt von einem dritten Untersucher zu einem Konsens-Rating zusammengeführt.

Bei übereinstimmender Beurteilung des Vorhandenseins bzw. Nichtvorhandenseins des jeweiligen Kriteriums wurden die angegebenen Sicherheitsscores der beiden Erst-Rater gemittelt, wobei der Mittelwert jeweils in Richtung auf das niedrigere Sicherheitsmass abgerundet wurde. Bei diskrepanter Kriterienbeurteilung überprüfte der Konsens-Rater erneut das Krankenblatt und führte eine Entscheidung herbei. In diesem Fall wurde das niedrigste Maß an Beurteilungssicherheit angenommen.

2.5. Datenverarbeitung

Sämtliche Item- und Kriteriumsbeurteilungen wurden in einem ablochfähigen Belegbogen dokumentiert.

Jedem einzelnen Patienten wurde zunächst eine dreistellige Code-Nummer zugewiesen, die Kodierbögen wurden durch Entfernen der Namen und Aufnahmedaten anonymisiert und die Namenslisten vernichtet.

Die auf den Belegbögen festgehaltenen Informationen wurden sodann auf Lochkarten übertragen und in die Datenbank der Honeywell-Bull-Großrechenanlage des Rechenzentrums der Universität Mainz eingelesen.

Nach Überspielen dieser Rohdaten auf handelsübliche Disketten-datenträger erfolgten zusätzliche Auswertungsschritte auf einem Personalcomputer (IBM-XT und Schneider-PC 1512HD20). Hierfür waren zuvor Anwendungsprogramme in der Programmiersprache Basic erstellt worden.

2.5.1. Statistische Verfahren

Die statistische Auswertung erfolgte durch Erstellen von (2x6) - Feldertafeln. Dieser Schritt wurde auf dem Großrechner des Universitätsrechenzentrums mit der Programmsammlung SPSS-X durchgeführt.

Alle Mehrfelder-Tafeln, die durch Beibehaltung der subjektiven Sicherheitsmasszahl entstanden waren, wurden in einfache Vierfeldertafeln überführt. 1-3 bedeutete hierbei ja die Abwesenheit des betreffenden Kriteriums, 4-6 dessen Anwesenheit. Wir setzten die Abwesenheit =0 und die Anwesenheit =1, so dass jetzt 0/1 skalierte Nominaldaten vorlagen.

Bei der weiteren Datenauswertung kamen ausschliesslich nichtparametrische Methoden wie Kappa, Sensitivität, Spezifität, prozentuale Übereinstimmung und positiv prädiktiver Wert (ppW 50) zum Einsatz. Alle werden weiter unten noch ausführlicher dargestellt und definiert werden.

2.5.1.1. Statistische Methodik zur Erfassung der Reliabilität

Das Konstrukt der Reliabilität ist ein Mass der Übereinstimmung, mit welcher unterschiedliche Untersucher Unterschiedliches unterscheiden können.

Die Reliabilität von Diagnosen hängt ab von Kriteriumsvarianz, Informationsausmass und der Beobachtungsvarianz (Spitzer und Fleiss 1975). Nach Meinung von Spitzer und Fleiss stellt die Kriteriumsvarianz den Hauptgrund für eine niedrige Reliabilität dar.

Daher sollte nach Meinung dieser und weiterer Autoren (u.a. Helzer 1977) Kriterienvarianz durch Anwendung expliziter diagnostischer Kriterien in Form von operationalisierten Diagnosesystemen vermindert werden.

Beobachtungsvarianz kann durch gleichzeitige Untersuchung eines Patienten durch zwei oder mehr Rater eliminiert werden. Dem Faktor der Informationsvarianz kann man durch Verwenden strukturierter Interviewverfahren begegnen. Im Falle der vorliegenden Untersuchung wurde versucht die Kriterienvarianz durch Anwendung operationaler Diagnostik, die Beobachtungsvarianz durch Teilnahme mehrerer Rater und die Informationsvarianz durch das jedem Rater zur Verfügung stehende normierte Krankenblatt zu vermindern.

Inter-Rater-Reliabilität stellt einen Sonderfall der allgemeinen Reliabilität dar. Bei der Untersuchung der Inter-Rater-Reliabilität sind die Ergebnisse zweier Rater, die das gleiche Krankenblatt ausgewertet hatten, zueinander in Bezug gesetzt worden.

Es interessiert also bei diesem spezielleren Gesichtspunkt die Frage, ob verschiedene Untersucher zu gleichen oder doch zumindest vergleichbaren Ergebnissen kommen. In diesem Fall ist die Reliabilität hoch einzustufen. Bei diskrepanten Ergebnissen wäre die Reliabilität niedrig.

2.5.1.1.2. Berechnung des Kappa-Koeffizienten

Als mathematisch-statistische Methoden zur objektiven Bestimmung der Inter-Rater-Reliabilität wurden prozentuale Übereinstimmung und der Kappa-Koeffizient berechnet. Kappa ist ein Maß der zufallskorrigierten Übereinstimmung. Es wurde von R. Cohen 1960 eingeführt und gilt heute als Standardverfahren bei kategorialen Variablen (Spitzer et al. 1967, 1979; Fleiss 1971; Mombour 1972; Spitzer und Fleiss 1974; Bech und Clemmesen 1983).

Die Formel des Kappa-Koeffizienten k lautet:

$$k = \frac{P_o - P_c}{1 - P_c}$$

Diese Formel setzt das Maß der beobachteten Übereinstimmung P_o in Beziehung zu der zufallserwarteten Übereinstimmung P_c , wobei sich P_c aus den Randsummen der $x \cdot x$ Feldertafeln errechnet.

In dieser Untersuchung sind alle Mehr-Feldertafeln auf 2×2 Feldertafeln reduziert, so daß sich das Problem der Gewichtung der einzelnen Felderbesetzungen nicht stellt (Everitt, 1968).

Kappa kann einen Wert zwischen den Extrempolen +1 und -1 annehmen. Entspricht die tatsächliche Übereinstimmung lediglich der Zufallserwartung, so wird $\kappa = 0$. Bei Werten von $\kappa > 0$ übersteigt die Übereinstimmung also die Zufallserwartung, Werte von $\kappa < 0$ zeigen ein Übereinstimmungsniveau an, das niedriger als die Zufallserwartung liegt.

Bei sehr niedrigen und sehr hohen Häufigkeitswerten (base-rate) unterschätzt Kappa die Übereinstimmung. Dieses Problem tritt auf, wenn einzelne Diagnosen oder Items Häufigkeiten von $< 20\%$ oder $> 80\%$ annehmen (Spitznagel und Helzer, 1985).

Um dem zu entgehen ist analog der Empfehlung einiger Autoren (Maurer et al. 1984; Kuny et al. 1982; Woggon et al. 1978, 1979) auch in der vorliegenden Arbeit auf die Berechnung von Kappa verzichtet worden, wenn die base-rates unterhalb von 10% lagen; der umgekehrte Fall (base-rates $> 90\%$) trat in unserem Datenmaterial nicht auf.

2.5.1.2. Untersuchung der Validität

Für die Validitätsbestimmung wurde ermittelt, in welchem Maße das auf seine Validität hin zu untersuchende Verfahren – hier das Krankenblatt-Rating – in der Lage ist, das Validierungskriterium – hier das Ergebnis des Interview-Ratings – zu schätzen. Maßzahlen für die Schätzgenauigkeit sind Assoziation (ermittelt durch den Chi-Quadrat-Vierfeldertest), Sensitivität (Prozentsatz der richtig erfassten Kriteriumsträger), Spezifität (Prozentsatz der richtig erfassten kriteriumsfreien Personen) und positiver prädiktiver Wert ppW 50 (Prozentsatz der richtig zugeordneten Fälle im Fall einer base-rate des Validierungskriteriums von 50%).

Alle genannten statistischen Berechnungen ergeben sich in ihrer mathematischen Formel unmittelbar aus der Besetzung der Vierfelder-Tafeln. Daher soll an dieser Stelle die Felderdefinition tabellarisch veranschaulicht werden.

Tabelle 6) Felderdefinition der verwendeten Vier-Felder-Tafeln

Validierungskriterien

	0	1
0	a	b
1	c	d

0 = nicht vorhanden

1 = vorhanden

Testkriterien

2.5.1.2.1. Berechnung der Sensitivität

Unter Sensitivität versteht man die Fähigkeit eines Kriteriums, die betroffenen Kriterienträger des Untersuchungsgutes richtig zu identifizieren. Bei einer Vierfeldertafel der angegebenen Felderdefinition errechnet sie sich als Quotient aus dem rechten, unteren Feld dividiert durch die Summe beider rechten Felder; das so erhaltene Ergebnis wird mit 100 multipliziert. Das Resultat drückt einen Prozentsatz aus. Formelhaft zusammengefasst ergibt sich also der folgende Rechengang:

$$\text{Sensitivität (ss)} = \frac{d}{b + d} \times 100\%$$

2.5.1.2.2. Berechnung der Spezifität

Spezifität bezeichnet die Fähigkeit, Nicht-Kriterienträger richtig zu erfassen. Auch die Spezifität stellt einen Quotienten dar, der aus der Besetzung einer Vierfeldertafel ableitbar ist. Die Formel lautet:

$$\text{Spezifität (sp)} = \frac{a}{a + c} \times 100\%$$

In diesem Falle werden in den Rechengang also die beiden linken Felder der Vierfeldertafel einbezogen.

2.5.1.2.3. Der Positiv Prädiktive Wert (ppW 50)

Für die Berechnung des positiv prädiktiven Wertes (ppW 50) wird zunächst die Besetzung der beiden rechten Felder b und d unter Beibehaltung ihrer Relation so zu b' und d' verändert, daß ihre Randsumme genauso groß wird, wie die Randsumme von a und c, also des Validierungskriteriums der base-rate von 50% erhält.

Sodann wird die Trefferquote in den Feldern a und d aufsummiert und der prozentuale Anteil der Gesamtstichprobe angegeben. Die mathematische Formel hierzu lautet:

$$\text{ppW 50} = \frac{a + d'}{a + b' + c + d'} \times 100$$

3. Ergebnisse

3.1. Reliabilität

Bei der Untersuchung der Inter-Rater-Reliabilität werden die Ergebnisse zweier Untersucher, die das gleiche Krankenblatt ausgewertet haben, zueinander in Beziehung gesetzt. Hierzu werden die prozentuale Übereinstimmung und der Kappa-Koeffizient berechnet (vgl. 2.5.1.1.).

Es sollen jeweils Ergebnisse auf Diagnosen-Ebene - also erstens die Major Depression in RDC, DSM-III und FDC, zweitens das Syndrom der endogenen Depression in RDC, DSM-III, FDC und der beiden Newcastle-Skalen (NC1 und NC2) und drittens die übrigen depressiven Syndrome - tabellarisch zusammengefaßt werden. Ein gleiches wird für die manischen Diagnosen erfolgen.

3.1.1. Reliabilität auf Diagnosen-Ebene

Zur Interpretation von Kappa haben Bech und Clemmesen (1983) folgende Einteilung vorgeschlagen. Danach sind Werte $\geq .80$ als exzellente Übereinstimmung zu werten, Werte zwischen $k .60$ und $.79$ zeigen eine gute Übereinstimmung. Nur als ausreichend ist Kappa zwischen $.40$ und $.59$ zu interpretieren, während bei Werten von $k \leq .40$ eine unzureichende Übereinstimmung angezeigt wird. Nach dieser Einteilung erscheint die Fähigkeit der hier untersuchten OPD, ein ausgeprägtes depressives Syndrom zu erfassen, in etwa identisch. Die detaillierten Ergebnisse hierzu sind der nachfolgenden Tabelle zu entnehmen.

Tabelle 1) Inter-Rater-Reliabilität der untersuchten OPD
1. Major Depression

Diagnosen	prozentuale Übereinstimmung	zufallskorrigierte Übereinstimmung (Kappa)
RDC Major Depressive		
Disorder (MDD)	89.9	.77
wiederkehrend-unipolare MDD	85.9	.59
psychotische MDD	88.9	.36
inaktivierende MDD	73.7	.46
agitierte MDD	93.9	.63
gehemmte MDD	76.8	.31
situative MDD	76.8	.24
einfache MDD	81.8	.62
bipolar 1	96.0	.64
bipolar 2	94.9	.68
DSM Major Depressive		
Episode (MDE)	89.9	.78
einzelne MDE	83.8	.51
wiederholte MDE	86.9	.62
bipolar, depressiv	99.0	.93*
FDC Depression	88.9	.77
primäre Depression	90.9	.81

* ungewichtetes Kappa;
Felderbestzung <10%

Die Kappa-Werte für Major Depressive Disorder (RDC), Major Depressive Episode (DSM) und Depression in FDC sind annähernd gleich und liegen nach der soeben vorgestellten Einteilung im guten Bereich.

Dies gilt jedoch nur für die Major Depression; in den verschiedenen Untergruppierungen finden sich besonders bei RDC überwiegend nur ausreichende bis unzureichende Werte.

Im Vergleich der Ergebnisse der Tabelle 2) liegen die berechneten Werte wiederum im gleichen Bereich guter zufallskorrigierter Übereinstimmung. Einzig der Wert für NC2 ist niedriger und kann nur als ausreichend beurteilt werden. Das niedrige Kappa von .33 bei DSM-III unterschätzt die tatsächliche Übereinstimmung, da in diesem Falle eine sehr niedrige base-rate von <10% vorliegt (siehe auch Helzer und Spitznagel 1985).

Tabelle 2) Inter-Rater-Reliabilität der untersuchten OPD
2. Endogene Depression

Diagnosen	prozentuale Übereinstimmung	zufallskorrigierte Übereinstimmung (Kappa)
RDC endogene MDD	87.9	.67
DSM MDE with melancholia	87.9	.33*
FDC Depression	88.9	.77
NC1	89.9	.66
NC2	87.9	.46

* ungewichtetes Kappa;
Felderbestzung <10%

Tabelle 3) Inter-Rater-Reliabilität der untersuchten OPD
3. andere depressive Störungen

Diagnosen	prozentuale Übereinstimmung	zufallskorrigierte Übereinstimmung (Kappa)
RDC Minor Depressive Disorder	90.9	.58
Minor DD mit Angst	86.9	.66
Intermittierende depressive Erkrankung	94.9	.52
DSM Dysthymie	97.0	.78
Anpassungsstörung	91.9	.30

Wiederum zeigen sich in dieser letzten Gruppe mit depressiven Syndromen (Tabelle 3) durchgehend befriedigende bis gute Werte der zufallskorrigierten Übereinstimmung mit Ausnahme der DSM-III-Kategorie "Anpassungsstörung mit depressiven Zügen".

In der Gruppe der manischen Syndrome (vgl. Tabelle 4) kann ebenfalls von einer sehr guten diagnostischen Übereinstimmung der Inter-Rater-Reliabilität gesprochen werden, wobei der Kappa-Koeffizient für das manische Syndrom bei RDC einen ausgezeichneten Wert zeigt. Das DSM-III zeigt von allen drei erfassten OPD die niedrigste Übereinstimmung. Zwei Kategorien (DSM-III: bipolar 2 und RDC: Hypomanie) zeigen eine nur niedrige Übereinstimmung.

Tabelle 4) Inter-Rater-Reliabilität der untersuchten OPD
4. Manie

Diagnosen	prozentuale Übereinstimmung	zufallskorrigierte Übereinstimmung (Kappa)
RDC Manie	94.9	.81
Hypomanie	98.0	.49
DSM manische Episode	94.9	.67
bipolar II	97.0	.65
FDC Manie	92.9	.74
primäre Manie	96.0	.77

3.1.2. Reliabilität auf Symptom-Ebene

3.1.2.1. Depressive Erkrankung

Aus den Ergebnissen der Untersuchungen zur Inter-Rater-Reliabilität auf Diagnosen-Ebene war eine gute bis zufriedenstellende Übereinstimmung ablesbar. Wie allerdings bereits in der Einleitung (1.1.) festgestellt wurde, sind Zweifel angebracht, ob angesichts teilweise recht differenzierter und oft nur unter Schwierigkeiten beurteilbarer Einzelitems, die in den verschiedenen operationalisierten Diagnosesystemen (OPD) zur Diagnosestellung beitragen bzw. obligat gefordert werden, noch eine genügend gute Übereinstimmung zwischen zwei Ratern besteht, die bei der Erstellung ihrer OPD-Diagnosen auf die Auswertung von Krankenblatttexten angewiesen sind, da unter Umständen ein Teil der Beobachtungen, die zur klinischen Diagnose führten, nicht im Krankenblatt niedergelegt worden sind.

Als Beispiele für solche Einzel-Items können die in verschiedenen OPD oft auch verschieden definierten Kriterien der Schlafstörungen, des Gewichtsverlustes, der qualitativ unterschiedlichen Art der depressiven Verstimmung bei endogener Depression, das Früherwachen, objektivierbare Hemmung oder Agitation und einige mehr angeführt werden.

Um diese sensiblen Parameter besser erfassen und bezüglich ihrer Reproduzierbarkeit bei verschiedenen Ratern beurteilen zu können, sollen in synoptischen Übersichten die jeweils geforderten Einzelsymptome der verschiedenen Diagnosegruppen tabellarisch erfasst und zusammengestellt werden.

Dazu wird zunächst in tabellarischer Form dargestellt, welche Items welchen OPD gemeinsam sind und worin sich die untersuchten OPD unterscheiden.

Tabelle1) Synopse der zur Diagnosestellung erforderlichen
Items in fünf konkurrierenden OPD

1. Major Depression und endogene Depression

Kriterienliste	RDC	DSM-III	FDC	NC1	NC2
1. Querschnittskriterien:					
abgrenzbare Periode dysphorischer					
Gestimmtheit (depressiv, traurig)	X	X	X		
besondere Qualität d. Depression	X	X		X	
fehlende Reaktivität	X	X			XX
Morgentief	X	X			XX
Verlust v. Vergnügen		X			
Früherwachen	X	XX			
Appetit-/Gewichtsänderung	X	X	X	X	
Schlafstörungen	X	X	X		
Energieverlust/Müdigkeit	X	X	X		
objektivierbare Veränderung der					
Psychomotorik	X	XX	(X)	X	XX
Interesseverlust	X	XX	X		
Selbstvorwürfe/Schuldgefühle	X	X	X	X	
Konzentrationsmangel	X	X	X		
Gedanken an Tod/Suizid	X	X	X		
Angst				X	
nihilistischer Wahn				X	X
Beschuldigen anderer				X	
2. Zeitkriterien:					
Dauer 2 Wochen	X	(X)			
Dauer 1 Monat			X		XX
plötzliches Auftreten					X
3. Schweregradkriterien:					
suchte um Behandlung nach	X				
Gleichbleibender Zustand					X
4. Ausschlußkriterien:					
Schizophrene Symptomatik	X	X			
Organische Ursache	X	X			
Psychologische Belastung		X		X	X
5. Anamnestiche Daten:					
abweichende Persönlichkeit				X	X
frühere Depression				X	

Anmerkungen:

XX = Symptom wird strenger
definiert

(X) = Symptom wird gefordert
ohne jede Definition

Die Übersicht zeigt deutlich, daß zwischen den beiden Diagnosesystemen RDC und DSM-III eine weitgehende Übereinstimmung auf Kriterienebene herrscht, wobei drei Symptome (Appetit- und Gewichtsänderung, Psychomotorik und Interesseverlust) im DSM-III strenger definiert werden. Zusätzlich wird die Kriterienliste um ein Einschluß- (Verlust von Vergnügen) und ein Ausschlußkriterium (psychologische Belastung) ergänzt. Die Feighner-Kriterien (FDC) fordern zur Diagnosestellung einer nicht näher bezeichneten Depression wesentlich weniger Symptome, was ihre Spezifität erheblich mindert (Nelson et al. 1981) und die damit erfaßten Patientengruppen unhomogen gestaltet. Die beiden Newcastle-Skalen erfassen zusätzlich noch Symptome der Anamnese, die bei allen anderen OPD hier fehlen, und differenzieren die mit den anderen OPD gemeinsamen Items stärker, so daß zwischen den Optionen "vorhanden", "wahrscheinlich" und "nicht vorhanden" gewählt werden kann. Sie verzichten allerdings dagegen auf Kriterien der Dauer der Erkrankung und auf Ausschlußkriterien, die sich auf mögliche organische Ursachen des depressiven Syndroms beziehen.

Tabelle 2) Synoptischer Vergleich der Inter-Rater-Reliabilität
von drei konkurrierenden OPD
1. Major Depression

Kriterien	prozentuale Übereinstimmung			kappa		
	RDC	DSM	FDC	RDC	DSM	FDC
depressive Verstimmung	84.8	86.0	88.9	.68	.66	.75
Appetit-/Gewichtsänderung	90.9	79.9	87.9	.79	.67	.71
Schlafstörungen	87.9	87.9	87.9	.75	.75	.75
Energieverlust/Müdigkeit	92.9	90.7	91.9	.85	.81	.83
objektivierbare Veränderung der Psychomotorik	77.8	76.9	79.8	.54	.47	.57
Interesseverlust	87.9	87.7	86.9	.75	.66	.73
Selbstvorwürfe/Schuldgefühle	91.9	92.1	90.9	.77	.67	.74
Konzentrationsmangel	89.9	88.4	89.9	.79	.69	.79
Gedanken an Tod/Suizid	88.9	84.0	88.9	.77	.75	.77
Dauer min. 1 Monat			84.8			.69
Mittelwerte:	88.1	84.9	87.8	.74	.68	.73

Tabelle 3) Synoptischer Vergleich der Inter-Rater-Reliabilität
von vier konkurrierenden OPD
1. endogene Depression

Kriterienliste	prozentuale Übereinstimmung				kappa			
	RDC	DSM	NC1	NC2	RDC	DSM	NC1	NC2
Besondere Qualität der Depression	81.8	78.8	84.5		.44	.37	.36	
fehlende Reaktivität	86.9	87.9		80.8	.47	.57		.57
Morgentief	88.9	89.9		94.9	.64	.67		.79
Interesseverlust	80.8	83.8			.58	.66		
Schuldgefühle	90.9	91.9	90.0		.73	.74	.68	
Früherwachen	85.9	91.9		94.9	.70	.59		.71
veränderte Psychomotorik	78.8	82.8	70.7	76.8	.56	.62	.31	.31
Appetitmangel	87.9	87.9			.72	.69		
Gewichtsverlust	86.9		91.9		.50		.59	
Interesseverlust/Verlust an Freude (RDC)	87.9				.75			
abweichende Persönlichkeit			91.9				.55	
psychol. Belastung			87.9	85.9			.66	.60
frühere Depression			86.9				.64	
Angst			85.9				.45	
nihilistischer Wahn			93.9	96.0			.53*	.72
Beschuldigen anderer			93.9				.22*	
plötzliches Auftreten				84.8				.03*
Dauer der Manifestation				82.8				.66
Phobien				98.0				.49*
gleichbleibender Zustand				78.8				.27
Mittelwerte:	85.7	86.8	87.7	87.4	.61	.61	.55	.52

* ungewichtetes Kappa bei Felderbestzung
<= 10%

Bei der vergleichenden Betrachtung der in Tabelle 3) aufgelisteten Ergebnisse finden wir insbesondere bei dem zur Diagnosestellung in allen OPD sehr wichtigen Einzelitem "besondere Qualität der Depression" sehr schlechte Werte der zufallskorrigierten Übereinstimmung (Kappa), die annähernd gleich niedrig .36 (NC1), .37 (DSM) und .44 (RDC) liegen. Bestätigt wird dieser Befund von den mit dieser "besonderen Qualität" mittelbar in Zusammenhang stehenden Items wie "fehlende Reaktivität" (RDC, DSM, NC2), wo die Ergebnisse nur marginal besser sind (.47, .57 und .57) und somit nach Bech und Clemmesen (1983) ebenfalls nur in die Gruppe der ausreichenden Übereinstimmung fallen. Damit korrespondiert auch das noch schlechtere Kappa von .27 des NC2-Items "gleichbleibender Zustand".

Ein wirklich gutes Maß der Inter-Rater-Reliabilität findet sich letztlich nur in den Unterpunkten "Morgentief", "Schuldgefühle" und "Früherwachen". Hier liegen die Kappa-Werte im Bereich einer guten Übereinstimmung.

3.1.2.2. Übrige depressive Syndrome

Im folgenden finden sich noch Tafeln, aus denen die Inter-Rater-Reliabilität der übrigen depressiven Syndrome ablesbar ist. Es handelt sich dabei um die Kategorien der "Minor Depressive Disorder" nach RDC und der "Dysthymie" sowie der "Anpassungsstörung mit depressiven Zügen", beide nach DSM-III:

Tabelle 4) Inter-Rater-Reliabilität
Minor Depressive Disorder (RDC)

Kriterienliste		prozentuale Übereinstimmung	zufallskorrigierte Übereinstimmung (Kappa)
A	Depressivität	83.3	.66
B1	Appetitmangel/Gewichtsverlust	90.0	.79
B2	Schlafstörung	87.9	.75
B3	Energieverlust	91.9	.83
B4	veränderte Psychomotorik	78.8	.56
B5	Interesseverlust	88.9	.77
B6	Schuldgefühle	90.9	.74
B7	Konzentrationsstörungen	89.9	.79
B8	Suizidalität	89.9	.79
B9	nonverbale depr. Zeichen	78.8	.56
B10	pessimistische Einstellung	78.8	.50
B11	Grübeln	87.9	.72
B12	Insuffizienz	84.8	.62
B13	reizbar, klagsam	81.8	.51
B14	fordernd, anklammernd	85.9	.41
B15	Selbstmitleid	85.9	.25*
B16	Hypochondrie	87.9	.20*
Mittelwert:		86.1	.62

* ungewichtetes Kappa bei Felder-
besetzung < 10%

Bei diesen Untergruppen der Diagnosesysteme finden wir zusammengefaßt eine wesentlich bessere zufallskorrigierte Übereinstimmung als in den verschiedenen Gruppen der endogenen Depression mit Mittelwerten von 86,1% und Kappa = .62 im Falle der Minor Depressive Disorder. Lediglich die Einzelitems "veränderte Psychomotorik", "nonverbale depressive Zeichen", "pessimistische Einstellung", "reizbar, klagsam" und "fordernd, anklammernd"

weisen Kappa-Werte des nur ausreichenden Bereiches zwischen $k \geq .40$ und $k \leq .59$ auf. Die Symptome "Selbstmitleid" und "Hypochondrie" leiden an einer unausgewogenen Felderbesetzung, so daß die hier errechneten Werte aus der Beurteilung ausgenommen sein sollen.

Für die Dysthymie nach DSM errechnen sich die folgenden Mittelwerte: prozentuale Übereinstimmung 84,9% und Kappa .63 . Analog der eben beschriebenen Diagnose "Minor Depressive Disorder" sind auch bei der "Dysthymie" die Kappa-Werte deutlich besser ausgefallen als bei der endogenen Depression. Die Einzel-Symptome "Reizbarkeit, sozialer Rückzug" und "mangelnde Reaktivität" zeigen eine geringere diagnostische Übereinstimmung.

Tabelle 5) Inter-Rater-Reliabilität
Dysthymie (DSM-III, 300.40)

Kriterienliste		prozentuale Übereinstimmung	zufallskorrigierte Übereinstimmung (Kappa)
A	Depressives Syndrom während der letzten 2 Jahre	98.0	.84
B	fluktuierende Manifestation	74.7	.50
C	depressive Stimmung/Interesse- verlust	85.9	.70
D1	Schlaflosigkeit/Hypersomnie	84.8	.69
D2	Energieverlust	90.9	.81
D3	Insuffizienzgefühl	81.8	.59
D4	nachlassende Leistung	90.9	.81
D5	Konzentrationseinbussen	84.8	.69
D6	sozialer Rückzug	78.8	.55
D7	Verlust von Freude	83.8	.67
D8	Reizbarkeit	82.8	.45
D9	Unfähig, Lob/Belohnung mit Vergnügen zu beantworten	83.8	.24
D10	Aktivitätsverlust	81.8	.63
D11	Pessimismus/Grüblen	83.8	.65
D12	Weinerlichkeit	83.8	.58
D13	Suizidalität	87.9	.75
Mittelwert		84.9	.63

In der Unterdiagnose des DSM "Anpassungsstörung mit depressiven Zügen" liegen die Ergebnisse allerdings deutlich schlechter. Hier finden sich Werte der prozentualen Übereinstimmung von 81% bzw. der zufallskorrigierten Übereinstimmung von $k=.47$.

Tabelle 6) Inter-Rater-Reliabilität
Anpassungsstörung (DSM-III, 309)

Kriterienliste	prozentuale Übereinstimmung	zufallskorrigierte Übereinstimmung (Kappa)
A Fehlangepasste Reaktion auf Belastung; Beginn innerhalb 3 Monate	83.3	.37
B Symptomatik übersteigt die zu erwartende Reaktion	77.8	.55
C keine Überreaktion auf Stress; keine Exacerbation einer psy- chischen Krankheit	73.7	.40
D Annahme: A verschwindet bei Wegfall der Belastung	89.9	.58
Mittelwert	.81.0	.47

3.1.2.3. Manische Erkrankung

Wir finden bei den Maniekriterien in allen drei verglichenen OPD beinahe identische Werte der prozentualen und zufallskorrigierten Übereinstimmung mit überwiegend guten Kappa-Werten. Lediglich in den Items "Ablenkbarkeit" und "exzessives Engagement" liegen die gefundenen Ergebnisse in einem unbefriedigenden Bereich. Die Operationalisierung der Diagnose Manie erscheint somit wesentlich reliabler reproduzierbar als die der depressiven Syndrome. Die einzelnen Ergebnisse können der nachfolgenden Tabelle entnommen werden.

Tabelle 7) Synoptischer Vergleich der Inter-Rater-Reliabilität von drei konkurrierenden OPD
2. manische Erkrankung

Kriterienliste	prozentuale Übereinstimmung			kappa		
	RDC	DSM	FDC	RDC	DSM	FDC
Periode(n) mit gehobener o. expansiver Stimmung	93.9	93.9	93.9	.80	.80	.80
Aktivitätssteigerung	93.7	93.9	93.9	.79	.79	.79
Logorrhoe	93.9	93.9	93.9	.79	.79	.79
Ideenflucht	92.9	90.8	93.9	.67	.71	.71
Gehobenes Selbstwertgef.	93.9	93.9	98.0	.69	.69	.89
Verminderter Schlaf	96.0	96.0	97.9	.82	.80	.89
Ablenkbarkeit	90.9	90.0	91.9	.35*	.36	.39
Exzessives Engagement in div. Aktivitäten	90.9	91.9		.42*	.51	
Sinnvolle Unterhaltung unmöglich	88.9			.30		
Schwere Beeinträchtigung im sozialen Kontext	91.9			.71		
Hospitalisierung	94.5			.79		
Dauer mind. 1 Woche	94.9			.83		
Dauer mind. 2 Wochen			92.9			.75
Keine schizophrene Sympt.	93.9	94.9		.71	.73	
keine Halluzinationen		90.9			.58	
kein bizarres Verhalten		90.9			.63	
nicht Folge einer organischen Psychose		91.9			.74	

* ungewichtetes Kappa bei Felderbesetzung <10%

Die Ergebnisse der Reliabilitätsstudie zusammenfassend bleibt festzuhalten, daß die diagnostischen Hauptkategorien aller hier untersuchten operationalen Diagnosesysteme (OPD) reliabel von verschiedenen Untersuchern reproduzierbar sind. Dies bezieht sich im einzelnen auf die Major Depressive Disorder nach RDC, die Major Depressive Episode nach DSM-III, die Depression nach FDC und auf die Gruppe der manischen Erkrankungen aller drei genannten OPD. Die prozentuale Übereinstimmung hierbei ist $\geq 90\%$, die Kappa-Werte liegen - orientiert man sich an der von Bech und Clemmesen (1983) empfohlenen Einteilung - in einem guten bis ausgezeichneten Bereich.

In den verschiedenen diagnostischen Untergruppen nimmt in beinahe allen Fällen Kappa niedrigere Werte an, während die prozentuale Übereinstimmung zumeist auf dem guten Niveau der Hauptkategorien verbleibt. Dieser Befund gilt auch für die Diagnose der endogenen Depression.

Auf der Symptom-Ebene ist ebenfalls ein zweigeteiltes Ergebnis zu finden. Hier sind die depressiven "Kernsymptome", wie sie in den Einzelitems der Major Depression festgelegt sind, also depressive Verstimmung, Schlafstörungen, Verlust des energetischen Potentials und Suizidalität mit guter Reliabilität aus den Aufzeichnungen der Krankenakte reproduzierbar.

In den verschiedenen Kriterienlisten für die Diagnose einer endogenen Depression der untersuchten OPD zeigt sich jedoch ein etwas anderes Bild.

Besonders die für die Endogenität eines depressiven Syndroms als charakteristisch beschriebenen Symptome der besonderen Qualität, der fehlenden Reaktivität, einer veränderten Psychomotorik und des Gewichtsverlustes nimmt Kappa Werte zwischen max. .57 und min. .31 an. Somit ist die zufallskorrigierte Übereinstimmung dieser Items nur noch ausreichend oder gar unzureichend.

3.2. Übereinstimmungsvalidität

Wie exemplarisch bereits in der Einführung (Abschnitt 1.2.) gezeigt wurde, sind die Definitionen der einzelnen diagnostischen Kategorien innerhalb der verschiedenen operationalisierten Diagnosesysteme durchaus unterschiedlich aufgebaut und zeigen eine verschieden ausgeprägte Restriktivität durch die unterschiedlich geforderte Anzahl zu erfüllender Items, wobei teilweise einzelne Kriterien noch genauer spezifiziert werden.

Zieht man zusätzlich die Möglichkeit in Betracht, daß nicht alle vielleicht tatsächlich vorhandenen Symptome auch im Krankenblatt dokumentiert wurden - besonders dann, wenn sie nicht schon beim Aufnahmebefund nachweisbar waren, sondern erst im stationären Verlauf hinzugetreten sind - so müssen sich Zweifel einstellen, ob eine Diagnosestellung durch die Auswertung von Krankengeschichten hinreichend valide sein kann.

Zur Klärung dieser Fragestellung wurde in der vorliegenden Arbeit das diagnostische Ergebnis der Krankenblattauswertung zu dem des strukturierten Aufnahmeinterviews, das mit der Present State Examination (PSE; Wing 1974) erhoben wurde, in Beziehung gesetzt. Gleichzeitig versuchten wir festzustellen, ob das Validitätskriterium durch einen zusätzlichen dritten, den sogenannten "Konsens-Rater" (vgl. 2.4.) besser erreicht werden kann.

Die folgenden Tabellen geben die erhaltenen Erkenntnisse wieder.

Tabelle 1) Übereinstimmungsvalidität der untersuchten OPD
1. Major Depression

Diagnosen	ss	sp	ppW ₅₀
	Erstrater (Konsensrater)		
RDC Major Depressive Disorder (MDD)	80.0(78.6)	82.2(80.0)	77.5(79.2)
wiederkehrend-unipolare MDD	52.0(77.7)	83.3(84.4)	67.6(81.1)
psychotische MDD	23.1(0)	94.0(96.2)	55.6(50.0)
inaktivierende MDD	23.4(25.0)	90.0(90.5)	56.7(57.8)
agitierte MDD	40.0(25.0)	94.2(94.0)	57.2(59.5)
gehemmte MDD	52.4(57.1)	75.0(87.2)	65.6(72.1)
situative MDD	63.2(70.0)	82.0(84.1)	72.6(77.0)
einfache MDD	65.9(70.6)	79.2(78.4)	73.9(74.6)
bipolar 1	55.5(50.0)	97.7(100.)	76.6(75.0)
bipolar 2	37.5(100.)	98.8(98.1)	63.1(99.1)
DSM Major Depressive Episode (MDE)	92.1(94.1)	76.3(83.8)	84.2(88.9)
einzelne MDE	54.5(61.5)	74.6(80.5)	65.0(71.0)
wiederholte MDE	55.0(77.7)	81.8(82.2)	68.4(80.0)
bipolar, depressiv	87.5(66.6)	94.4(98.0)	90.9(82.3)
DC Depression	77.7(66.6)	75.0(75.0)	76.4(70.8)
primäre Depression	65.9(68.4)	81.1(77.1)	73.5(72.7)
Mittelwerte:	57.8(61.8)	84.9(86.8)	70.3(74.4)

Bei der Interpretation der Ergebnisse, wie sie sich in Tabelle 1) darstellen, ist als wichtigstes Resultat festzuhalten, daß vor allem die großen Diagnosen-Gruppen wie Major Depressive Disorder (MDD; RDC), Major Depressive Episode (MDE; DSM-III) und die Depression nach den Feighner-Kriterien valide reproduzierbar erscheinen. Sie alle weisen in den vorgenommenen statistischen Prüfungen bezüglich Sensitivität, Spezifität und positiv prädiktivem Wert gute bis ausgezeichnete Werte aus. Die jeweiligen Chi-Quadrat Signifikanzgrenzen erreichen ein Niveau von $p \leq 0.001$.

Im Falle der differenzierteren Unterkategorien aller operationalen Diagnosesysteme finden sich jedoch durchgehend schlechtere Werte.

Während die errechneten Werte der Spezifität, also der Fähigkeit, Nicht-Kriterienträger zu erfassen und auszuschließen, noch ein sehr befriedigendes Niveau aufweisen, sinkt vor allen Dingen die Sensitivität ab. Das heißt mit anderen Worten, es können in diesen Fällen positive Kriterienträger nicht mehr ausreichend sicher identifiziert werden. Am deutlichsten ist dieser Befund in den Untergruppen der RDC nachweisbar. Hier befindet sich bei der psychotischen MDD und der inaktivierenden MDD die Sensitivität unter der 30%-Marke. In gleicher Weise betroffen sind die beiden Arten der bipolaren Verlaufsformen. Parallel dazu sinken auch die Werte des ppW50. Bei der Beurteilung einer veränderten Psychomotorik ist ein vergleichbarer Trend zu beobachten. In den Untergruppen des DSM-III sind die Veränderungen weniger deutlich; vor allem die Gruppe "bipolar, depressiv" erscheint dabei valide und reliabel reproduzierbar.

Weiterhin kann festgehalten werden, daß die Einbeziehung eines sogenannten "Konsensraters" in beinahe allen Fällen die Reproduzierbarkeit von aus Krankengeschichten erhobenen Diagnosen auch in den problematischen Untergruppen, die eben angesprochen wurden, verbessern kann. Die hierfür errechneten und in Klammern angegebenen Werte liegen bei allen Sub-Diagnosen in einem vergleichbaren oder höheren Bereich.

Tabelle 2) Übereinstimmungsvalidität der untersuchten OPD
2. Syndrome der endogenen Depression

Diagnosen	ss	sp	ppW ₅₀
	Erstrater (Konsensrater)		
RDC endogene MDD	59.1(44.4)	85.3(91.1)	72.2(76.7)
DSM MDE with melancholia	11.1(33.3)	88.6(96.1)	49.8(64.7)
FDC Depression	77.7(66.6)	75.0(75.0)	76.4(70.8)
NC1	58.8(25.0)	86.2(92.0)	72.5(58.5)
NC2	38.4(16.6)	90.5(95.8)	64.5(56.3)
	49.0(37.2)	85.1(90.0)	67.1(65.4)

Tabelle 2) zeigt die Ergebnisse der Validitätsstudie für die Gruppe der endogenen Depression. Wie aufgrund der besonderen und restriktiveren Kriterien für diese Diagnose zu erwarten war, liegen die hier erzielten Ergebnisse unter denen der verschiedenen Grobkategorien der Major Depression.

Am deutlichsten ist dies aus den unbefriedigend niedrigen Werten der Sensitivität bei gleichzeitig hoher Spezifität ablesbar. Eine beträchtliche Anzahl Diagnosenträger kann durch die Auswertung der Krankengeschichte nicht mehr erfasst werden, während umgekehrt Nicht-Kriterienträger gut identifiziert und ausgeschlossen werden können.

Besonders gravierend scheint dies bei der DSM-III-Diagnose "Major Depressive Episode with melancholia" der Fall zu sein. Hier finden wir die niedrigsten Werte sowohl für Sensitivität als auch für den positiv prädiktiven Wert. Etwas befriedigender liegen die Verhältnisse bei der endogenen Depression nach RDC und bei der NC1-Skala. Eine Mittelstellung nimmt der NC2 ein.

Diese Ergebnisse lassen den Schluß zu, daß die Kriterien der "MDE with melancholia" wesentlich restriktiver und damit einhergehend diese Diagnose schlechter reproduzierbar ist als die der anderen OPD.

Wiederum kann dieses Resultat durch Einbeziehen eines Konsensraters verbessert werden, ohne dadurch jedoch das Ergebnis in vergleichbaren Maßen wie bei Tabelle 1) verbessern zu können, d.h. das Konsensrating erhöht zwar die Reliabilität, nicht aber die Validität der untersuchten Diagnosen.

Die verbleibenden Restkategorien der depressiven Syndrome (vgl. Tabelle 3) zeigen in der Ergebnisdokumentation - darin den in den anderen beiden Tabellen dargestellten Unter-Diagnosen vergleichbar - ebenfalls unbefriedigende Werte bezüglich Sensitivität und ppW 50. Erneut finden sich in den beiden hier angeführten DSM-III-Diagnosen die niedrigsten Werte.

Tabelle 3) Übereinstimmungsvalidität der untersuchten OPD
3. andere depressive Störungen

Diagnosen		ss	sp	ppW ₅₀
		----- Erstrater (Konsensrater)		
RDC	Minor Depressive Disorder	50.0(40.0)	85.5(84.1)	70.3(62.1)
	Minor DD mit Angst	36.0(30.8)	80.5(85.3)	58.3(58.1)
	Intermittierende depressive Erkrankung	--	--	--
DSM	Dysthymie	14.3(50.0)	96.6(96.1)	55.5(73.1)
	Anpassungsstörung	12.5(0)	98.9(98.0)	55.9(49.0)
Mittelwert:		28.2(30.2)	90.4(90.9)	60.0(60.5)

Anmerkung: Die diagnostische Kategorie der intermittierenden depressiven Erkrankung war im untersuchten Patientengut nicht besetzt.

Tabelle 4) Übereinstimmungsvalidität der untersuchten OPD
4. manische Syndrome

Diagnosen	ss	sp	ppw ₅₀
	-----	-----	-----
	Erstrater	(Konsensrater)	
RDC Manie	71.5(87.5)	100 (100)	85.7(93.8)
Hypomanie	100 (100)	96.1(100)	45.8(97.2)
DSM manische Episode	77.7(100)	96.6(97.9)	87.2(98.9)
FDC Manie	63.2(58.3)	98.7(100)	80.9(79.2)
primäre Manie	60.0(75.0)	97.7(97.8)	78.8(86.4)
Mittelwert:	74.5(84.2)	97.8(99.1)	75.7(91.1)

Die in Tabelle 4) dargestellten Ergebnisse für die manische Erkrankung bietet gleichmäßig für alle untersuchten operationalen Diagnosesysteme wiederum wie schon bei den diagnostischen Grobkategorien der Depression in Tabelle 1) durchgehend gute bis hervorragende Werte in allen statistisch berechneten Parametern. Erneut kann festgestellt werden, daß die Ergebnisse durch einen Konsensrater Verbesserungen zeigen; allerdings fallen diese bei der gut reproduzierbaren Diagnose einer Manie weniger deutlich ins Gewicht.

Zusammenfassend kann als Ergebnis der Validitätsstudie festgestellt werden, daß sich die Hauptdiagnosen eines ausgeprägten depressiven Syndroms (MDD nach RDC, MDE nach DSM-III und Depression nach FDC) ebenso wie die Diagnose einer manischen Erkrankung durch Auswertung von Krankenakten gut reproduzierbar erweisen.

Dies gilt aber nicht - oder nur in eingeschränkter Form - für die verschiedenen diagnostischen Untergruppen der Depression.

Durch einen zusätzlichen Konsensrater können alle Ergebnisse in ihren statistischen Parametern noch weiter verbessert werden, ohne daß jedoch für die genannten diagnostischen Untergruppen das gleiche befriedigende Niveau wie bei den Hauptkategorien erreichbar ist.

4. Diskussion

Die vorliegende Untersuchung stellt die Frage nach der Reliabilität und Validität der aufgrund einer Auswertung von Krankenblatttexten erhaltenen Diagnosen. Es interessieren dabei sowohl die Inter-Rater-Reliabilität, also ob verschiedene Untersucher zu vergleichbaren Ergebnissen gelangen, als auch die Validität, in diesem Fall also ob die Diagnose der Krankenblattauswertung mit der Diagnose eines Live-Ratings übereinstimmt.

4.1. Diskussion der Ergebnisse der Reliabilitätsprüfung

Die Ergebnisse der Reliabilitätsprüfung zusammenfassend bleibt festzuhalten, daß die diagnostischen Hauptkategorien aller hier untersuchten operationalen Diagnosesysteme (OPD) reliabel von verschiedenen Untersuchern reproduzierbar sind. Dies bezieht sich im einzelnen auf die Major Depressive Disorder nach RDC, die Major Depressive Episode nach DSM-III, die Depression nach FDC und auf die Gruppe der manischen Erkrankungen aller drei genannten OPD. Die prozentuale Übereinstimmung hierbei ist $\geq 90\%$, die Kappa-Werte liegen - orientiert man sich an der von Bech und Clemmesen (1983) empfohlenen Einteilung - in einem guten bis ausgezeichneten Bereich.

In den verschiedenen diagnostischen Untergruppen nimmt in beinahe allen Fällen Kappa niedrigere Werte an, während die prozentuale Übereinstimmung zumeist auf dem guten Niveau der Hauptkategorien verbleibt. Dieser Befund gilt auch für die Diagnose der endogenen Depression.

Auf der Symptom-Ebene ist ebenfalls ein zweigeteiltes Ergebnis zu finden. Hier sind die depressiven "Kernsymptome", wie sie in den Einzelitems der Major Depression festgelegt sind, also depressive Verstimmung, Schlafstörungen, Verlust des energetischen Potentials und Suizidalität mit guter Reliabilität aus dem Aufzeichnungen der Krankenakte reproduzierbar.

In den verschiedenen Kriterienlisten für die Diagnose einer endogenen Depression der untersuchten OPD zeigt sich jedoch ein etwas anderes Bild.

Besonders die für die Endogenität eines depressiven Syndroms als charakteristisch beschriebenen Symptome der besonderen Qualität, der fehlenden Reaktivität, einer veränderten Psychomotorik und des Gewichtsverlustes nimmt Kappa Werte zwischen max. .57 und min. .31 an. Somit ist die zufallskorrigierte Übereinstimmung dieser Items nur noch ausreichend oder gar unzureichend.

Dieses letztgenannte Phänomen war bereits bei Philipp und Maier (1987) nachweisbar: je weiter die diagnostische Hierarchie sich verzweigt, also z.B. im Schritt von der Major Depression zur endogenen Depression, desto niedriger werden die Werte für die zufallskorrigierte Übereinstimmung und für die diagnostische Assoziation. Die Inter-Rater-Reliabilität ist also anders gesagt umso größer, je weitgespannter die Klassifikation bleibt (Spitzer et al. 1975).

4.1.1. Die diagnostischen Hauptkategorien

Das in der vorliegenden Untersuchung gefundene Ergebnis, daß die unter 4.1. bereits genannten diagnostischen Grobkategorien der Major Depression und der Manie bei allen drei hier untersuchten operationalen Diagnosesystemen (OPD) mit hoher Reliabilität reproduzierbar sind, deckt sich

mit den Befunden zahlreicher anderer Autoren (Spitzer et al. 1978; Klerman, Endicott und Spitzer 1979; Andraesen et al. 1982; Wittchen et al. 1985; Fuhrer et al. 1986). Die genannten Autoren fanden ihre Ergebnisse allerdings nicht auf der Basis der Auswertung von Krankengeschichten, sondern hatten eine Interviewtechnik – entweder als joint-rating oder als test-retest-rating durchgeführt – benutzt.

Aus Gründen der besseren Übersichtlichkeit sollen die gefundenen Ergebnisse zunächst nochmals tabellarisch im Vergleich zu den Resultaten anderer Autoren dargestellt werden.

Tabelle 1) Zufallskorrigierte Übereinstimmung der diagnostischen Hauptkategorien (Kappa)

Diagnosen	RDC		DSM-III		FDC	
	-----		-----		-----	
Studien	MDD	Manie	MDE	Manie	Depr.	Manie
eigene Daten	.77	.81	.78	.67	.77	.74
Fuhrer 1986			.78			
Wittchen 1985	.74		.78			
Hyer 1982			.59			
Robins 1981	.68	.65	.63	.65	.66	.66
Spitzer 1979			.77	.70		
Spitzer 1978	.76	.85				
Helzer 1977					.55	.82
Spitzer 1975	.97	.89				

Die in der vorliegenden Arbeit gefundenen Resultate lassen wohl den Schluß zu, daß auch die retrospektive Krankenblattauswertung keinen Informationsverlust bedeutet, also die Reliabilität mit dieser Methode nicht sinkt. Der Einwand von Craig et al. (1982), daß einzelne Symptome - obwohl klinisch vorhanden - im Krankenblatt nicht festgehalten worden sein könnten, womit er die dort beobachteten Abweichungen zwischen klinischer DSM-II Diagnose und erhaltener DSM-III Diagnose nach Krankenblattauswertung zu erklären versucht, trifft also zumindest im Falle der hier diskutierten diagnostischen Hauptkategorien nicht zu. Auch Hyler et al. (1982) widersprechen der Ansicht, daß unterschiedlicher Informationsgehalt, also geringer Umfang des Krankenblattes, für ein schlechteres Ergebnis verantwortlich seien. Darüberhinaus ist aus obenstehender Tabelle klar ersichtlich, daß die aus Krankenblattauswertungen gewonnenen Daten hinsichtlich ihrer Werte der zufallskorrigierten Übereinstimmung denen mit Interview-Technik gewonnenen Daten nicht unterlegen sind.

Natürlich braucht hier nicht mehr besonders betont zu werden, daß es sich bei diesen operationalen Diagnosen der Major Depression um eine deutlich heterogene Gruppe depressiver Syndrome handelt. Dieser Aspekt ist mehrfach eindringlich von anderen Autoren hervorgehoben worden (Nelson und Charney 1980; Carroll et al. 1980).

Zum Problem der Konkordanz zwischen den Systemen, also zur Frage, ob in den verschiedenen OPD in den vergleichbaren diagnostischen Kategorien auch die gleichen Krankheitsbilder gemeint sind, fanden Singerman et al. (1981) im Vergleich DSM vs. RDC $k=1.0$, DSM vs. FDC $k=.89$ und FDC vs. RDC $k=.89$ bezüglich der Major Depression. Im Falle der Diagnose Manie waren Kappa 1.0 , $.96$ und $.96$ respektive. Er zog daraus den Schluß, daß die Diagnosen Depression und Manie in den drei OPD äquivalent seien. Im Gegensatz dazu fanden andere Untersucher (Philipp et al. 1985b) eine wesentlich ungünstigere diagnostische Übereinstimmung zwischen verschiedenen OPD. Allerdings beschränkte sich diese Untersuchung auf die Kategorie der endogenen Depression. In diesem Falle nahm Kappa Werte $\leq .5$ an.

4.1.2. Die diagnostischen Untergruppen

4.1.2.1. Die endogene Depression

Aufgrund der komplexeren Auffächerung der für die Diagnose einer MDD, endogenous type (RDC) geforderten Einzelkriterien sinkt die Inter-Rater-Reliabilität dieser Untergruppe. Wir fanden eine zufallskorrigierte Übereinstimmung von $k=.67$. Dies liegt zwar unter der ausgezeichneten Übereinstimmung der MDD ($k=.77$), liegt aber nach Bech und Clemmesen (1983) in einem als gut einzustufenden Bereich. Auch in den von Spitzer, Endicott und Robins (1978) publizierten Daten war die Inter-Rater-Reliabilität dieser Diagnose niedriger als die der übergeordneten Kategorie "Major Depressive Disorder". Hier waren in dem Teil der Untersuchung, der sich mit der Test-Retest-Methode beschäftigte, ein Kappa-Wert von .58 für die endogene Depression gefunden worden.

Bereits Nelson et al. (1981) stellten in ihrer Studie über die Melancholie-Kriterien des DSM-III fest, daß das Kriterium "besondere Qualität der depressiven Verstimmung" zwar als sehr aussagefähig hinsichtlich der Unterscheidung endogen - nicht-endogen beurteilt wird, aber nur selten als zutreffend beurteilt wurde; dieses Ergebnis wurde auch von anderer Seite bestätigt (Craig et al. 1982). Auch in der vorliegenden Untersuchung zeigt dieses Einzelitem die schlechtesten Werte (für RDC $k=.44$, für DSM $k=.37$ und für NC1 $k=.36$). Copolov et al. (1986) fordern daher eine klare und eindeutige Definition, was denn unter dem Begriff der "besonderen Qualität der depressiven Verstimmung" zu verstehen sei.

Die deutlichste prädiktive Aussagekraft zur Unterscheidung zw. endogen und nicht-endogen sehen Nelson et al. in den Einzelitems "mangelnde Reaktivität" und "veränderte Psychomotorik".

In der vorliegenden Untersuchung erreichen aber auch diese als besonders diskriminierend erachteten Items nur Werte einer ausreichenden zufallskorrigierten Übereinstimmung. Es sind dies für das Kriterium "mangelnde Reaktivität" $k=.47$ (RDC), $k=.57$ (DSM) und $k=.57$ (NC2) und für die "veränderte Psychomotorik" $k=.56$ (RDC), $k=.62$ (DSM), und $k=.31$ (NC1 und NC2).

Letztlich erscheinen also ausgerechnet die als gut zwischen endogener und nicht-endogener Depression diskriminierenden Einzelitems als am wenigsten reliabel reproduzierbar. Ein wirklich gutes Maß an Reliabilität finden wir nur bei den Kriterien "Morgentief", "Schuldgefühle" und "Früherwachen", die Kappa-Werte einer guten Übereinstimmung aufweisen (vgl. 3.1.2.1., Tabelle 3). Die Stellung der Diagnose einer endogenen Depression auf der Basis von Krankenblattauswertungen muß daher für alle hier beurteilten OPD sehr zurückhaltend bewertet werden. Einschränkend muß allerdings auch angemerkt werden, daß im vorliegenden Studien-Setting die genannten Symptome explizit in den Krankengeschichten dokumentiert sein mußten, da andernfalls das Symptom als abwesend beurteilt wurde, auch wenn es wahrscheinlich vorgelegen haben mag.

Zu diesem Problem meinen Behar et al. (1980), daß in einer retrospektiv angelegten Studie wie dieser, Häufigkeit und Intensität bestehender Symptome wahrscheinlich unterschätzt werden.

4.1.2.2. Übrige depressive Syndrome

Die weitere Aufgliederung der RDC-Unterdiagnosen einer Major Depressive Disorder (vgl. Abschnitt 3.1., Tabelle 1) zeigt ein Absinken der Reliabilität aller diagnostischen Unterkategorien. Werte der zufallskorrigierten Übereinstimmung (Kappa) variieren zwischen $k=.31$ (gehemmte MDD) und $k=.68$ (Bipolarität). Diese geringere diagnostische Reliabilität der RDC-Untergruppen wird auch an anderer Stelle berichtet (Spitzer et al. 1978; Andraesen et al. 1982).

Für die RDC-Diagnose einer "Minor Depressive Disorder" fanden wir eine prozentuale Übereinstimmung von 90.9% bei einem Kappa von .58. Spitzer et al. (1979) berichten für die gleiche Diagnose eine zufallskorrigierte Übereinstimmung von $k=.64$, was angesichts der Tatsache der dort angewandten joint-interview-Technik, die ohnehin bessere Werte der Inter-Rater-Reliabilität erwarten lässt (Hyler et al. 1982), für eine ausreichende Reproduzierbarkeit dieser Diagnosen durch Krankenblattauswertung spricht.

In der DSM-III Untergruppe "Dysthymie" fanden wir eine zufallskorrigierte Übereinstimmung von $k=.78$. Dies liegt im gleichen Bereich wie bei Wittchen et al. (1985), der zur Datenerhebung ein standardisiertes Interview (DIS) benutzte und ein Kappa von .81 fand. Fuhrer et al (1986) berichtet in der DSM-Untergruppe "Dysthymie" ein Kappa von .39 und im Falle der Cyclothymen Erkrankung sogar nur $k=.27$ (joint-interview-Technik), wobei diese beiden sehr niedrigen Werte wohl durch die Tatsache mitbestimmt worden waren, daß es sich um eine minimale Fallzahl handelte.

Insgesamt muß jedoch festgehalten werden, daß auch die recht niedrigen Werte der Reliabilitätsstudie für die diagnostischen Untergruppen immer noch höher liegen, als die von Spitzer und Fleiss (1974) errechneten Kappa-Werte ihrer Re-analyse früherer Reliabilitätsstudien. Sie liegen vielmehr in einem den mittels Interviewtechnik erhaltenen Ergebnissen vergleichbaren Bereich, was bei aller Einschränkung auf die prinzipielle Fähigkeit zur reliablen Reproduzierbarkeit dieser Diagnosen auch bei Krankenblattauswertungen schließen läßt.

4.2. Diskussion der Ergebnisse der Validitätsprüfung

4.2.1. Die diagnostischen Hauptkategorien

Erneut kann als Quintessenz der hier durchgeführten Untersuchung zur Validität - darin dem Ergebnis der Reliabilitätsstudie vergleichbar - festgehalten werden, daß die Hauptdiagnosen des depressiven Syndroms (Major Depression in RDC und DSM-III und Depression nach FDC) sowie die Diagnose einer manischen Erkrankung durch retrospektive Auswertung von Krankengeschichten mit guter Validität reproduzierbar sind.

Für die Diagnose der Major Depressive Disorder (MDD) nach RDC lagen Sensitivität (ss), Spezifität (sp) und Positiv Prädiktiver Wert (ppW 50) um 80% . Die Major Depressive Episode (MDE) nach DSM-III lag in ss und ppW 50 sogar noch höher, während hier die Spezifität auf einem Niveau knapp unter der 80%-Marke blieb. Geringfügig niedriger lagen die Resultate der FDC. In gleicher Größenordnung lagen auch die gefundenen Ergebnisse der manischen Erkrankung.

Alle Resultate konnten durch die Einbeziehung eines Konsensraters noch verbessert werden (vgl. Abschnitt 3.2., Tabelle 1).

Diese Daten decken sich im wesentlichen mit denen anderer Publikationen (Robins, Helzer et al. 1981), die in ihrer Studie zur Validierung des Diagnostic Interview Schedule (DIS) ein test-retest-Design zwischen Laien-Interviewern einerseits und Psychiater-Interviewern andererseits, das in diesem Fall das Validierungskriterium darstellt, verwendet hatten, wobei jedoch kritisch angemerkt werden sollte, ob das so konzipierte Studiensetting von Robins et al. nicht vielmehr eine Reliabilitätsprüfung als eine Validitätsprüfung darstellte.

Dagegen fanden andere Autoren (Helzer et al. 1978) in ihrer Nachbeurteilung der Krankengeschichten von 97 stationär behandelten Patienten eine deutliche Abweichung der erhaltenen Diagnosen im Vergleich zum strukturierten Aufnahmeinterview auch bei den Diagnosen Depression und Manie. Leider ist aber zu dieser eben genannten Studie keine absolute Vergleichbarkeit gegeben, da die Methodik zwar vergleichbar war, jedoch als statistische Maßzahl der Übereinstimmung der Kappa-Koeffizient berechnet wurde. Helzer fand für die Gruppe der Depressionen ein Kappa von .41, für die manischen Erkrankungen ein Kappa von .70 , wobei allerdings unter "Depressionen" in diesem Fall das gesamte Spektrum depressiver Syndrome subsummiert ist.

4.2.2. Die diagnostischen Untergruppen

Wiederum darin den Ergebnissen der Reliabilitätsstudie vergleichbar sinkt die Reproduzierbarkeit der einzelnen diagnostischen Untergruppen depressiver Syndrome durch Nachauswertung von Krankengeschichten in Bezug zum Aufnahmeinterview ab. Während die Spezifität hoch bleibt, nehmen Sensitivität und positiv prädiktiver Wert unbefriedigend niedrige Resultate an. Dies gilt insbesondere für die Untergruppen der Major Depressive Disorder der RDC, die die meisten Subspezifizierungen zulässt (vgl. Abschnitt 3.2., Tabelle 1 und Abschnitt 6., Tabellarischer Anhang). Einzige Ausnahme dieses Trends sind die bipolar-depressiven Verlaufsformen. Auch die Einbeziehung eines zusätzlichen Konsensraters verbessert die Ergebnisse kaum.

In der Gruppe der endogenen Depression zeigt ganz besonders die Diagnose "MDE with melancholia" (DSM-III), wie schwierig sie reproduzierbar erscheint. Bei extrem niedriger Sensitivität und hoher Spezifität liegt der positiv prädiktive Wert um 50% . Eine augenscheinlich große Zahl Diagnosenträger kann also nicht mehr angemessen identifiziert werden. Dies mag in erster Linie mit der hohen Restriktivität erklärbar sein, die diese Diagnose aufweist, und die in zwei obligaten Eingangskriterien und strenger als bei den RDC definierten Anschlußkriterien zu sehen ist. Daneben muß ebenfalls die Frage diskutiert werden, ob die Untergruppe "MDE with melancholia" nicht einfach nur Patienten mit einem besonders stark ausgeprägten depressiven Syndrom erfasst, also keine eigentliche nosologische Einheit darstellt (Zimmerman et al. 1986b) und deren konzeptionelle Validität von einem Teil der in den USA praktizierenden Psychiater angezweifelt wird (Jampala et al. 1986).

Etwas besser liegen die Verhältnisse für die Untergruppe der endogenen Depression in den beiden OPD-Systemen RDC und FDC. Bei einer Sensitivität um 60% liegen Spezifität und ppW 50 bei 85% bzw. 72% . Dieses im Vergleich zur entsprechenden DSM-III-Kategorie bessere Ergebnis ist mit der weitgefassteren, weniger restriktiven Definition dieser Klassifikationen zu erklären. Daß die Diagnose der endogenen Depression in diesen beiden OPD häufiger gestellt wird, wurde von anderen Autoren ebenfalls bestätigt (Davidson et al. 1984a; Philipp und Maier 1985a und 1986a; Zimmerman et al. 1986a). Die gleichen Autoren gelangen daher zu der Schlußfolgerung, daß die auch hier untersuchten OPD hinsichtlich ihrer diagnostischen Zuordnung eines gegebenen Patientengutes nur eingeschränkt vergleichbar seien.

Offensichtlich kommt es zu einem erheblichen Verlust von Einzelinformationen bei der Dokumentation des klinischen Bildes in der Krankengeschichte, der eine retrospektive Diagnosestellung erschwert.

Auch die noch verbleibenden Diagnosen depressiver Syndrome, also die "Minor Depressive Disorder" (RDC) und die beiden DSM-III-Kategorien "Dysthymie" und "Anpassungsstörung", bleiben unter den Ergebnissen der diagnostischen Hauptgruppen - teilweise sogar unter denen der endogenen Depression, was wiederum in besonderem Maße für die Diagnosen des DSM-III zutrifft.

Dieses Ergebnis wird sicherlich auch durch die Tatsache mitbestimmt, daß im vorliegenden Untersuchungsgut die absolute Fallzahl der diese Kriterien erfüllenden Patienten sehr niedrig ist.

5. Zusammenfassung

Thema der vorliegenden Arbeit war die Untersuchung von Inter-Rater-Reliabilität und Validität operationalisierter Diagnostik auf der Basis von Krankenblattauswertungen.

Es wurden die Krankengeschichten von $n=97$ Patienten ausgewertet und nach den Kriterienlisten für affektive Erkrankungen der operationalisierten Diagnosesysteme (OPD) Research Diagnostic Criteria (RDC), Diagnostic and Statistical Manual of Mental Diseases (DSM-III), Feighner-Criteria (FDC) und der beiden Newcastle-Skalen (NC1 und NC2) beurteilt. Alle 97 Patienten waren bereits zum jeweiligen Aufnahmedatum mittels der Present State Examination (PSE) standardisiert exploriert worden. Die damals gewonnenen Daten waren gleichfalls nach den genannten OPD beurteilt worden.

Neben dem Referenten waren an der Auswertung der Krankengeschichten insgesamt vier weitere Assistenzärzte der Psychiatrischen Klinik der Universität Mainz beteiligt. Ein weiterer Kollege führte ein sogenanntes "Konsensrating" durch (vgl. Abschnitt 2.1. und 2.4.)

Im Falle der Reliabilitätsprüfung wurden die Ergebnisse zweier Rater, die das gleiche Krankenblatt ausgewertet hatten, zueinander in Beziehung gesetzt. In der statistischen Analyse der Ergebnisse wurden prozentuale und zufallskorrigierte Übereinstimmung (Kappa-Statistik; vgl. Abschnitt 2.5.1.) berechnet. Dabei zeigten die diagnostischen Grobkategorien der Major Depression und der Manie eine ausgezeichnete zufallskorrigierte Übereinstimmung (Kappa) zw. den beiden Ratern für alle hier untersuchten OPD. Für die Major Depression fanden sich Werte von $k=.77$ (RDC), $k=.78$ (DSM-III) und $k=.77$ (FDC).

Die entsprechende statistische Analyse der Diagnose "Manie" zeigte für RDC ein Kappa von .81, bei DSM-III war $k=.67$ und bei FDC zeigte Kappa einen Wert von .77 .

In den differenzierteren Untergruppierungen der OPD hingegen lagen die gefundenen Resultate deutlich niedriger und lagen in einem nur ausreichenden ($.40 \leq k \leq .59$) oder sogar unzureichenden Bereich ($k \leq .39$). Dies gilt leider auch für die diagnostische Kategorie der endogenen Depression. Hier zeigte Kappa für RDC bzw. NC1 einen Wert von .67 bzw. .66; die entsprechenden DSM-III-Diagnose dagegen war mit einem Wert von $k=.33$ nicht mehr ausreichend reliabel reproduzierbar.

Zur Überprüfung der Validität war das Ergebnis des Krankenblattratings mit dem eines Interview-Ratings verglichen worden. Zusätzlich sollte durch Einbeziehung eines Konsensraters festgestellt werden, ob die gefundenen Resultate so nochmals in ihrer Validität verbessert werden können.

Die statistische Analyse erfolgte durch Berechnung von Sensitivität (ss), Spezifität (sp) und positiv prädiktivem Wert (ppW_{50}). Wiederum konnten die großen Diagnosegruppen (Major Depression und Manie) mit hoher Validität reproduziert werden. Die Einbeziehung eines Konsensraters verbesserte das Ergebnis zusätzlich (vgl. Abschnitt 3.2.). In den diagnostischen Untergruppen dagegen sinken alle statistischen Parameter deutlich ab, wobei besonders ss und ppW_{50} betroffen sind. Der Konsensrater kann diese Tendenz nicht in dem erhofften Ausmaß korrigieren.

Es kann also abschließend betont werden, daß die diagnostischen Hauptkategorien der Major Depressive Disorder (RDC), der Major Depressive Episode (DSM-III) und der Depression nach FDC ebenso wie die einzelnen Diagnosen der manischen Erkrankung aller drei genannten operationalisierten Diagnosesysteme (OPD) ausreichend valide und reliabel durch Auswertung von Krankenblatttexten reproduzierbar sind.

Dies gilt jedoch nur in einem sehr eingeschränkten Maße für die verschiedenen diagnostischen Untergruppen der genannten OPD einschließlich der jeweiligen Diagnose für die endogene Depression, da die teilweise sehr differenzierten Kriterien sich nicht mehr mit der gewünschten Reliabilität und Validität aus den Aufzeichnungen der Krankengeschichten reproduzieren lassen. Allerdings fanden auch zahlreiche frühere Untersucher ein Absinken der Reliabilität und Validität dieser Unter-Diagnosen.

Die Einbeziehung eines Konsensraters erbringt nochmals eine geringe Validitätssteigerung, die jedoch nicht genügend hoch ausfällt, um die genannten Diagnosegruppen auf das aussagefähige Niveau der diagnostischen Hauptkategorien anzuheben.

6. Tabellarischer Anhang

Im tabellarischen Anhang sollen aus Gründen besserer Übersichtlichkeit die differenzierten Kriterien zur Diagnosestellung der in der vorliegenden Arbeit benutzten operationalisierten Diagnosen

Research Diagnostic Criteria (RDC)

Diagnostic and Statistical Manual of Mental Diseases (DSM-III)

Feighner Diagnostic Criteria (FDC)

im Wortlaut der deutschsprachigen Editionen aufgelistet werden. Auf die entsprechende Nummerierung der Tabellen wurde im Text der Dissertation verwiesen, soweit dies notwendig war.

Weiterhin findet sich im tabellarischen Anhang ein Dokumentationsbogen, wie er zur Datenerhebung für die vorliegende Arbeit benutzt worden war.

Tabelle 1) Operationalisierte Diagnosekriterien
Research Diagnostic Criteria (RDC)
Major Depressive Disorder (MDD)

A. Eine oder mehrere voneinander abgrenzbare Perioden mit dysphorischer Gestimmtheit oder überzeugendem Verlust an Interessen oder Freude.

B. Mindestens fünf der folgenden Symptome müssen als Teil einer Erkrankung vorhanden sein, um die Erkrankung als "sicher" vorhanden zu bewerten.

1. Mangelnder Appetit oder Gewichtsverlust oder Appetitzunahme oder Gewichtszunahme (Änderungen von 1/2kg pro Woche oder mehreren Wochen oder 5 kg pro Jahr, wenn keine besondere Diät eingehalten wird).
 2. Schlafstörungen oder Zunahme der Schlafdauer.
 3. Energieverlust, Schwächegefühl oder Müdigkeit.
 4. Psychomotorische Agitiertheit oder Verlangsamung (aber nicht nur das bloße subjektive Gefühl der Unruhe oder Gehemmtheit).
 5. Verlust an Interessen oder Spaß an üblichen Aktivitäten, einschließlich sozialer Kontakte oder Sexualität (wenn dies auf eine Periode begrenzt ist, in der der Patient Wahnideen hatte oder halluzinierte, sollten diese Symptome nicht mit einbezogen werden).
 6. Selbstvorwürfe oder ausgeprägte oder inadäquate Schuldgefühle (beides kann auch wahnhaft sein).
 7. Klagen oder Hinweise auf eine verminderte Fähigkeit zu denken oder sich zu konzentrieren, wie z.B. Denkverlangsamung oder Entschlußlosigkeit (falls diese Symptome mit deutlichen formalen Denkstörungen einhergehen, sollten sie hier nicht berücksichtigt werden).
 8. Wiederkehrende Gedanken an den Tod oder an Suicid oder irgendwelche Suicidhandlungen.
-

Tabelle 2) Operationalisierte Diagnosekriterien
RDC, Major Depressive Disorder, endogenous type

I und II müssen erfüllt sein:

- I) Die Kriterien einer Major Depressive Disorder (MDD) sind erfüllt
 - II) Mindestens ein Kriterium der Gruppe A) und mindestens sechs aus A) und B) sind erfüllt
-

- A1) Besondere Qualität der depressiven Verstimmung, d.h. die depressive Gestimmtheit wird als etwas eindeutig anderes erlebt als jene Stimmung, die nach dem Tode einer geliebten Bezugsperson erlebt wird
 - A2) Fehlende Reaktivität auf die Änderung der Umgebung, d.h. während der Depression fühlt sich der Patient auch dann nicht einmal zeitweise besser, wenn sich etwas Erfreuliches ereignet
 - A3) Die Stimmung ist regelmäßig morgens schlechter
 - A4) Durchgängiger Verlust von Interesse oder Vergnügen

 - B1) Selbstvorwürfe oder ausgeprägte oder inadäquate Schuldgefühle
 - B2) Frühes morgendliches Erwachen oder Durchschlafstörungen
 - B3) Objektivierbare psychomotorische Verlangsamung oder Agitation
 - B4) Appetitmangel
 - B5) Gewichtsverlust ($\geq 1\text{kg/Woche}$ über mehrere Wochen oder $\geq 10\text{kg/1 Jahr}$)
 - B6) Interesseverlust oder Verlust an Vergnügen im Bereich üblicher Aktivitäten oder vermindertes sexuelles Interesse
-

**Tabelle 3) Operationalisierte Diagnosekriterien
RDC, Minor Depressive Disorder**

Diese Kategorie ist für nicht-psychotische Episoden oder Perioden einer Erkrankung vorbehalten, bei der die ausgeprägteste Störung eine relativ anhaltende depressive Gestimmtheit ist, ohne daß dies dem Vollbild eines depressiven Syndroms entspräche, wie es unter Major Depression Disorder charakterisiert ist. Sie soll ebenso von "intermittierenden depressiven Erkrankungen" unterschieden werden.

A Eine Krankheitsepisode, bei der eine relativ persistierende depressive Gestimmtheit das klinische Zustandsbild dominiert (oder eine gleichwertige Ängstlichkeit vorhanden ist). Die depressive Gestimmtheit kann als depressiv, traurig, mutlos, hoffungslos, niedergeschlagen oder "erledigt" beschrieben werden.

B. Zwei oder mehrere der unten aufgeführten Symptome sind als Teil der Erkrankung in Erscheinung getreten:

1. Appetitmangel oder Gewichtsverlust oder Appetitzunahme oder Gewichtszunahme (eine Änderung von einem 1/2 kg pro Woche über mehrere Wochen oder 5 kg über den Zeitraum von einem Jahr ohne besondere Diät).
2. Schlafstörungen oder Zunahme der Schlafdauer.
3. Energieverlust, Schwächegefühle oder Müdigkeit.
4. Psychomotorische Agitiertheit oder Verlangsamung (aber nicht bloß das subjektive Gefühl der Ruhelosigkeit oder Gehemmtheit).
5. Interesseverlust oder Verlust an Vergnügen bei den üblichen Aktivitäten einschließlich sozialer Kontakte oder Sexualität (diese Symptome sollten nicht berücksichtigt werden, wenn sie auf eine Zeitperiode beschränkt bleiben, in der der Patient wahnhaft war oder halluzinierte).
6. Selbstvorwürfe oder ausgeprägte oder inadäquate Schuldgefühle (beides kann auch wahnhaft sein).
7. Klagen oder Hinweise von verminderter Fähigkeit zu denken oder sich zu konzentrieren, wie z.B. verlangsamtes Denken oder Entschlußlosigkeit (diese Symptome sollten hier nicht berücksichtigt werden, wenn sie zusammen mit offensichtlichen formalen Denkstörungen auftreten).
8. Immer wiederkehrende Gedanken an Tod oder Selbstmord, oder irgendwelche suicidale Handlungen.
9. Nichtverbale Manifestationen einer Depression wie z.B. Weinerlichkeit oder trauriger Gesichtsausdruck.
10. Pessimistische Einstellung
11. Grübeln über vergangene od. gegenwärtige unangenehme Ereignisse.
12. Beherrschtsein von Insuffizienzgefühlen.
13. Übelnehmend, reizbar, ärgerlich oder klagsam.
14. Fordernd oder anklammernd.
15. Selbstmitleid.
16. Ausgeprägtes Besorgtsein um die körperliche Befindlichkeit.

Tabelle 4) Operationalisierte Diagnosekriterien
Diagnostic and Statistical Manual of Mental Disorders
(DSM-III). Major Depressive Episode (MDE)

A bis E müssen erfüllt sein

A. Dysphorische Stimmung oder Verlust von Interesse oder Vergnügen im Bereich aller oder fast aller gewohnten Aktivitäten und Freizeitbeschäftigungen. Die dysphorische Stimmung ist durch Symptome wie die in folgenden genannten gekennzeichnet: depressiv, traurig, bedrückt, hoffnungslos, niedergeschlagen, reizbar. Die Stimmungsstörung muß prominent und relativ anhaltend, jedoch nicht notwendigerweise das dominierende Symptom sein. Nicht eingeschlossen sind vorübergehende Verschiebungen von einer anderen Qualität dysphorischer Verstimmung, z.B. von Angst zu Depressivität oder zu Ärger, wie man es etwa in Zuständen akuter psychotischer Floridität sehen kann.

B. Mindestens vier der folgenden Symptome sind jedes für sich fast jeden Tag über einen Zeitraum von mindestens zwei Wochen vorhanden gewesen:

- 1) schlechter Appetit oder deutlicher Gewichtsverlust (ohne Diätbedingungen) oder Steigerung von Appetit und Gewicht;
- 2) Schlaflosigkeit oder Hypersomnie;
- 3) psychomotorische Agitiertheit oder Hemmung (jedoch nicht das bloß subjektive Gefühl von innerer Unruhe oder Verlangsamung);
- 4) Verlust von Interesse oder Vergnügen im Bereich der gewohnten Aktivitäten, oder Abnahme der Libido (jedoch nur dann, wenn diese auch ausserhalb von Verlaufsabschnitten nachweisbar sind, die durch Wahn oder Halluzinationen gekennzeichnet sind);
- 5) Verlust von Energie; Müdigkeit;
- 6) Gefühl der Wertlosigkeit, Selbstvorwürfe, oder exzessive oder unangemessene Schuldgefühle; (jedes von diesen kann auch wahnhaft sein);
- 7) Klagen über oder Objektivierbarkeit von Denk- und Konzentrationsstörungen, wie etwa Denkverlangsamung oder Entschlußlosigkeit, soweit diese nicht im Zusammenhang mit einer ausgeprägten Lockerung der Assoziationen (Zerfahrenheit) oder Inkohärenz stehen;
- 8) wiederkehrende Todesgedanken, Suizidvorstellungen, Todeswünsche oder Suizidversuch.

- C. Keines der folgenden Symptome dominiert das klinische Bild ausserhalb des unter A und B gekennzeichneten depressiven Symptoms, d. h. vor seinem Einsetzen oder nach seiner Remission:
- 1) Beschäftigtsein mit stimmungsinkongruentem Wahn oder Halluzinationen;
 - 2) bizarres Verhalten.
- D. Das depressive Syndrom ist nicht einer Schizophrenie, einer schizophreniformen Psychose, oder einer paranoiden Störung überlagert.
- E. Das depressive Syndrom ist nicht die Folge einer organischen Hirnerkrankung bzw -Funktionsstörung, und nicht die Folge eines unkomplizierten Verlusterlebnisses.
-

Tabelle 5) Operationalisierte Diagnosekriterien
DSM-III, MDE with melancholia

I und II müssen erfüllt sein

I) Die Kriterien einer MDE sind erfüllt

II) A1 und A2, sowie mindestens drei Kriterien von
B1 bis B6 treffen zu

A1) Verlust des Vergnügens im Bereich aller oder fast
aller Aktivitäten

A2) Fehlende Reaktivität auf normalerweise vergnügliche
Ereignisse, fühlt sich bei vergnüglichen Ereignissen
nicht einmal vorübergehend besser

B1) Besondere Qualität der depressiven Verstimmung,
d.h. die depressive Gestimmtheit wird als etwas
eindeutig anderes erlebt als jene Stimmung, die
nach dem Tod einer geliebten Bezugsperson erlebt wird

B2) Die Stimmung ist morgens regelmäßig schlechter

B3) Früherwachen von mind. 2 Stunden vor der gewohnten
Aufwachzeit

B4) Ausgeprägte psychomotorische Hemmung oder Agitation

B5) Gravierende Appetitlosigkeit oder Gewichtsminderung

B6) Exzessive oder unangemessene Schuldgefühle

Tabelle 6) Operationalisierte Diagnosekriterien
DSM-III, Dysthymie

A. Während der letzten zwei Jahre (bei Kindern und Heranwachsenden einem Jahr) hat der Betreffende fast die gesamte Zeit unter Symptomen zu leiden gehabt, die für das depressive Syndrom charakteristisch sind, die jedoch nicht schwer und lang anhaltend genug waren um die Kriterien für Major Depressive Disorder zu erfüllen.

B. Die Manifestationen des depressiven Syndroms können relativ persistierend oder aber von Perioden normaler Stimmung über einen Zeitraum von Tagen bis zu mehreren Wochen unterbrochen sein, nicht jedoch länger als jeweils ein paar Monate.

C. Während der depressiven Periode besteht entweder eine prominente depressive Stimmung (d. h. traurig, bedrückt, "im Keller"), oder aber ein ausgeprägter Verlust von Interesse und Vergnügen in allen oder fast allen gewohnten Aktivitäts- u. Freizeitbereichen.

D. Während der depressiven Periode sind mindestens drei der folgenden Symptome vorhanden:

1. Schlaflosigkeit oder Hypersomnie;
2. niedriges Energieniveau oder chronische Müdigkeit;
3. Gefühle der Unzulänglichkeit, Verlust der Selbstachtung oder Selbstentwertung;
4. Nachlassen der Leistungsfähigkeit in Schule, Arbeit oder Haushalt;
5. Nachlassen von Aufmerksamkeit, Konzentration oder Denkfähigkeit;
6. sozialer Rückzug;
7. Verlust des Interesse oder der Freude an vergnüglichen Aktivitäten;
8. Reizbarkeit oder exzessiver Ärger (bei Kindern gegenüber Eltern oder Erziehern);
9. Unfähigkeit Lob oder Belohnung mit erkennbarem Vergnügen zu beantworten;
10. weniger aktiv oder gesprächiger als normal, oder Gefühl verlangsamt oder ruhelos zu sein;
11. pessimistische Einstellung zur Zukunft, Grübeln über Dinge der Vergangenheit oder Gefühle des Selbstmitleids;
12. Weinerlichkeit oder Jammrigkeit;
13. Wiederholte Gedanken an Tod oder Suizid.

E. Fehlen psychotischer Symptome wie Wahn, Halluzinationen, Fehlen von Inkohärenz und Lockerung der Assoziationen.

F. Wenn sich die Störung einer vorbestehenden psychischen Erkrankung überlagert (z.B. Zwangskrankheit, Alkoholabusus) kann die depressive Stimmung dennoch eindeutig bezüglich Intensivität und Funktionsbeeinträchtigung von der gewohnten Stimmungslage des betreffenden unterschieden werden.

Tabelle 7) Operationalisierte Diagnosekriterien
Feighner Diagnostic Criteria (FDC)
1. Depression

A bis C müssen erfüllt sein.

A. Verstimmung, die durch Eigenschaftsworte wie die folgenden charakterisiert werden kann: depressiv, traurig, melancholisch, versagen, hoffungslos, niedergeschlagen, reizbar, ängstlich, besorgt oder entmutigt.

B. Mindestens fünf der folgenden Kriterien sind für eine definitive, vier für eine wahrscheinliche Diagnosenstellung erforderlich:

1. schlechter Appetit oder Gewichtsverlust von 1 kg pro Woche oder 10 kg in einem Jahr ohne Diätbedingungen.
2. Schlafstörungen (Schlaflosigkeit oder Hypersomnie).
3. Energieverlust, Erschöpfbarkeit, Müdigkeit.
4. Agitiertheit oder Hemmung.
5. Interesseverlust im Bereich der gewohnten Aktivität oder Libidoeinbuße.
6. Selbstvorwürfe oder Schuldgefühle (beide können wahnhaft sein).
7. Klagen über oder Objektivierbarkeit von Denk- oder Konzentrationsstörungen, wie etwa Denkverlangsamung oder Durcheinanderlaufen der Gedanken.
8. Wiederkehrende Gedanken an Tod oder Suizid einschließlich von Gedanken am liebsten tot sein zu wollen.

C. Dauer der Erkrankung mindestens einen Monat.

6.2. Kommentar zum Belegbogen

Auf den folgenden vier Seiten findet sich ein Exemplar des Belegbogens, wie er für die vorliegende Erhebung verwendet worden war.

Dabei findet sich links oben ein Kodierfeld, das Patient, Rating und Untersucher identifiziert. Darunter beginnt der eigentliche Erhebungsteil.

Jede der drei Spalten ist in sich nochmals gedrittelt, wobei in der ersten Spalte die Bezeichnung des jeweils zu beurteilenden Items ablesbar ist. Also z.B. RDC 9A als Bezeichnung der Research Diagnostic Criteria, Major Depressive Disorder. 9B₁₋₃ setzt die Kriterienliste weiter fort (vgl. Tabellarischer Anhang, Tabelle 1). Dies gilt analog für alle weiteren Diagnosen.

Die zweite Spalte bezeichnet mit fortlaufenden Nummern die Position der entsprechenden Items aus der ersten Spalte auf den Lochkarten, auf die die Informationen vor Einlesen in die Großrechenanlage des Rechenzentrums der Universität Mainz, übertragen worden waren.

Die dritte Spalte beinhaltet die eigentliche Symptom-Beurteilung der Items aus Spalte 1). Wie bereits in Abschnitt 2.4. erläutert wurde, meinen dabei die Ziffern 1-3 die Abwesenheit des Symptoms, die Ziffern 4-6 dessen Anwesenheit. Diese Einteilung bietet zugleich die Möglichkeit, die subjektive Sicherheit bei der Beurteilung zu dokumentieren. Die Ziffern 1 und 6 wären die Extrempole einer jeweils sehr hohen Beurteilungssicherheit, die Ziffern 3 und 4 sagen aus, daß der Rater sich bei der Beurteilung, ob das Symptom vorgelegen hat, unsicher war.

Neue Diagnosen und neue Lochkarten werden in der ersten Spalte angezeigt. Die Kartenfolge 28 bis 33 erfasst die Symptom-Ebene der vorliegenden Untersuchung, während die Karte 34 die gefundenen Diagnosen zusammenfaßt.

ame: _____

Untersucher: _____

1-3			
4-5			)
6-7	9		)
8-9	2	8	

depressives Syndrom

10-13				
-------	--	--	--	--

C					
19	1	2	3	4	5 6
20	1	2	3	4	5 6
21	1	2	3	4	5 6
22	1	2	3	4	5 6
23	1	2	3	4	5 6
24	1	2	3	4	5 6
25	1	2	3	4	5 6
26	1	2	3	4	5 6
27	1	2	3	4	5 6
28		0		1	2
29	1	2	3	4	5 6
30	1	2	3	4	5 6
31	1	2	3	4	5 6
32	1	2	3	4	5 6
34	1	2	3	4	5 6
35		0		1	2 3
36	1	2	3	4	5 6
37		0		1	2
38	1	2	3	4	5 6

9fA1	39	1	2	3	4	5	6
9fA2	40	1	2	3	4	5	6
9fA3	41	1	2	3	4	5	6
9fA4	42	1	2	3	4	5	6
9fB1	43	1	2	3	4	5	6
9fB2	44	1	2	3	4	5	6
9fB3	45	1	2	3	4	5	6
9fB4	46	1	2	3	4	5	6
9fB5	47	1	2	3	4	5	6
9fB6	48	1	2	3	4	5	6
9g1	50	1	2	3	4	5	6
9g2	51	1	2	3	4	5	6
9g3	52	1	2	3	4	5	6
9g4	53	1	2	3	4	5	6
9g5	54	1	2	3	4	5	6
9g6	55	1	2	3	4	5	6
9h1	57	1	2	3	4	5	6
9h2	58	1	2	3	4	5	6
9h3	59	1	2	3	4	5	6
9h4	60	1	2	3	4	5	6
9h5	61	1	2	3	4	5	6
9i	62				0	1	2
9i'	63	1	2	3	4	5	6
9j	64	1	2	3	4	5	6
9k	65	1	2	3	4	5	6 7
10A	66	1	2	3	4	5	6
10B1	67	1	2	3	4	5	6
10B2	68	1	2	3	4	5	6
10B3	69	1	2	3	4	5	6
10B4	70	1	2	3	4	5	6

10B5	71	1	2	3	4	5	6
10B6	72	1	2	3	4	5	6
10B7	73	1	2	3	4	5	6
10B8	74	1	2	3	4	5	6
10B9	75	1	2	3	4	5	6
10B10	76	1	2	3	4	5	6
10B11	77	1	2	3	4	5	6
10B12	78	1	2	3	4	5	6
10B13	79	1	2	3	4	5	6
10B14	80	1	2	3	4	5	6
Karte	8-9	2	9				
10B15	10	1	2	3	4	5	6
10B16	11	1	2	3	4	5	6
10C	12		0		1	2	
10C'	13	1	2	3	4	5	6
10D	14	1	2	3	4	5	6
10E	15	1	2	3	4	5	6
10F	16	1	2	3	4	5	6
10a	17		0		1	2	
10a'	18	1	2	3	4	5	6
11A	19	1	2	3	4	5	6
11B	20	1	2	3	4	5	6
11C	21	1	2	3	4	5	6
11D	22	1	2	3	4	5	6
11E	23	1	2	3	4	5	6
11F	24	1	2	3	4	5	6
7	25		0		1	2	3
8	26		0		1	2	3
7+8	27	1	2	3	4	5	6
4A1	28	1	2	3	4	5	6
4A2	29	1	2	3	4	5	6
4A3	30	1	2	3	4	5	6
4B	32	1	2	3	4	5	6
4C	33	1	2	3	4	5	6

DD

35	1	2	3	4	5	6
36	1	2	3	4	5	6
37	1	2	3	4	5	6
38	1	2	3	4	5	6
39	1	2	3	4	5	6
40	1	2	3	4	5	6
41	1	2	3	4	5	6
42	1	2	3	4	5	6
43	1	2	3	4	5	6
.	.	.	.	.	.	.
45	1	2	3	4	5	6
46	1	2	3	4	5	6
47	1	2	3	4	5	6
.	.	.	.	.	.	.
49	1	2	3	4	5	6
50	1	2	3	4	5	6
51	1	2	3	4	5	6
3	.	.	.	.	.	.
52	1	2	3	4	5	6
53	1	2	3	4	5	6
54	1	2	3	4	5	6
.	.	.	.	.	.	.
56	1	2	3	4	5	6
57	1	2	3	4	5	6
58	1	2	3	4	5	6
59	1	2	3	4	5	6
60	1	2	3	4	5	6
61	1	2	3	4	5	6

300.4

A	64	1	2	3	4	5	6
B	65	1	2	3	4	5	6
C	66	1	2	3	4	5	6
D1	67	1	2	3	4	5	6
D2	68	1	2	3	4	5	6
D3	69	1	2	3	4	5	6
D4	70	1	2	3	4	5	6
D5	71	1	2	3	4	5	6
D6	72	1	2	3	4	5	6
D7	73	1	2	3	4	5	6
D8	74	1	2	3	4	5	6
D9	75	1	2	3	4	5	6
D10	76	1	2	3	4	5	6
D11	77	1	2	3	4	5	6
D12	78	1	2	3	4	5	6
D13	79	1	2	3	4	5	6

Karte 8-9

310

E	11	1	2	3	4	5	6
F	12	1	2	3	4	5	6

309.xx

A	14	1	2	3	4	5	6
B	15	1	2	3	4	5	6
C	16	1	2	3	4	5	6
D	17	1	2	3	4	5	6
E	18	1	2	3	4	5	6

296.82

21	1	2	3	4	5	6
----	---	---	---	---	---	---

FDC

1

1A	22	1	2	3	4	5	6
1B1	23	1	2	3	4	5	6
1B2	24	1	2	3	4	5	6
1B3	25	1	2	3	4	5	6
1B4	26	1	2	3	4	5	6
1B5	27	1	2	3	4	5	6
1B6	28	1	2	3	4	5	6
1B7	29	1	2	3	4	5	6
1B8	30	1	2	3	4	5	6

1C

1.1.

B

NCI

1	53	0	1	2
2	54	0	1	2
3	55	0	1	2
4	56	0	1	2
5	57	0	1	2
6	58	0	1	2
7	59	0	1	2
8	60	0	1	2
9	61	0	1	2
10	62	0	1	2

NCII

1	63	0	1	2
2	64	0	1	2
3	65	0	1	2
4	66	0	1	2
5	67	0	1	2
6	68	0	1	2
7	69	0	1	2
8	70	0	1	2
9	71	0	1	2
10	72	0	1	2

Kask Ende

isches Syndrom

3-9 **31**

10	1 2 3 . 4 5 6
11	1 2 3 . 4 5 6
12	1 2 3 . 4 5 6
13	1 2 3 . 4 5 6
14	1 2 3 . 4 5 6
15	1 2 3 . 4 5 6
16	1 2 3 . 4 5 6
17	1 2 3 . 4 5 6
18	1 2 3 . 4 5 6
19	1 2 3 . 4 5 6
20	1 2 3 . 4 5 6
21	1 2 3 . 4 5 6
22	1 2 3 . 4 5 6
23	1 2 3 . 4 5 6
24	1 2 3 . 4 5 6
25	1 2 3 . 4 5 6
26	1 2 3 . 4 5 6
27	1 2 3 . 4 5 6
28	1 2 3 . 4 5 6
29	1 2 3 . 4 5 6

I

31	1 2 3 . 4 5 6
32	1 2 3 . 4 5 6
33	1 2 3 . 4 5 6
34	1 2 3 . 4 5 6
35	1 2 3 . 4 5 6
36	1 2 3 . 4 5 6
37	1 2 3 . 4 5 6
38	1 2 3 . 4 5 6
39	1 2 3 . 4 5 6
40	1 2 3 . 4 5 6
41	1 2 3 . 4 5 6

D	43	1 2 3 . 4 5 6
E	44	1 2 3 . 4 5 6
296.0	47	1 2 3 . 4 5 6
296.1	48	1 2 3 . 4 5 6
296.4	49	1 2 3 . 4 5 6
296.6	50	1 2 3 . 4 5 6
301.13	51	1 2 3 . 4 5 6
A	51	1 2 3 . 4 5 6
B	52	1 2 3 . 4 5 6
C	53	1 2 3 . 4 5 6
D	54	1 2 3 . 4 5 6
E	55	1 2 3 . 4 5 6
296.7	57	1 2 3 . 4 5 6
FDL	58	1 2 3 . 4 5 6
A	58	1 2 3 . 4 5 6
B1	59	1 2 3 . 4 5 6
B2	60	1 2 3 . 4 5 6
B3	61	1 2 3 . 4 5 6
B4	62	1 2 3 . 4 5 6
B5	63	1 2 3 . 4 5 6
B6	64	1 2 3 . 4 5 6
C	66	1 2 3 . 4 5 6
1.1.	68	1 2 3 . 4 5 6

Karte Ende Schizophrenes Syndrom

Karte 8-9 **32**

1A1	10	1 2 3 . 4 5 6
1A2	11	1 2 3 . 4 5 6
1A3	12	1 2 3 . 4 5 6
1A4	13	1 2 3 . 4 5 6
1A5	14	1 2 3 . 4 5 6
1A6	15	1 2 3 . 4 5 6
1A7	16	1 2 3 . 4 5 6
1A8	17	1 2 3 . 4 5 6

1B	19	1 2 3 . 4 5 6
1C	20	1 2 3 . 4 5 6
1.1.	22	1 2 3 . 4 5 6
1.2.	23	1 2 3 . 4 5 6
1.3.	24	1 2 3 . 4 5 6
1.4.	25	1 2 3 . 4 5 6
1.5.	26	1 2 3 . 4 5 6
DSM III		
295.1	27	1 2 3 . 4 5 6
A1	27	1 2 3 . 4 5 6
A2	28	1 2 3 . 4 5 6
A3	29	1 2 3 . 4 5 6
A4	30	1 2 3 . 4 5 6
A5	31	1 2 3 . 4 5 6
A6	32	1 2 3 . 4 5 6
B	34	1 2 3 . 4 5 6
C	35	1 2 3 . 4 5 6
D	36	1 2 3 . 4 5 6
E	37	1 2 3 . 4 5 6
F	38	1 2 3 . 4 5 6
295.1	40	1 2 3 . 4 5 6
295.2	41	1 2 3 . 4 5 6
295.3	42	1 2 3 . 4 5 6
295.9	43	1 2 3 . 4 5 6
295.6	44	1 2 3 . 4 5 6
295.4	45	1 2 3 . 4 5 6
298.8		
A	46	1 2 3 . 4 5 6
B1	47	1 2 3 . 4 5 6
B2	48	1 2 3 . 4 5 6
B3	49	1 2 3 . 4 5 6
B4	50	1 2 3 . 4 5 6
C	52	1 2 3 . 4 5 6
D	53	1 2 3 . 4 5 6
E	54	1 2 3 . 4 5 6

			Karte 18-91			
			Depr. Syndrom			
			RDC 9	0 1 2		
			9b	11-12		
			9c	13	0 1	
			9d	14	0 1 2	
			9e	15	0 1	
			9f	16	0 1 2	
			9g	17	0 1	
			9h	18	0 1	
			9i	19	0 1 2	
			9j	20	0 1	
			10	21	0 1 2	
			10a	22	0 1 2	
			11	23	0 1 2	
			7	24	0 1 2 3	
			8	25	0 1 2 3	
			4	26	0 1	
			DSM 296x	27	0 1	
			296.2	28	0 1	
			296.3	29	0 1	
			296.5	30	0 1	
			296.x3	31	0 1	
			300.4	32	0 1	
			309.x	33	0 1	
			FDC 1.	34	0 1	
			1.1.	35	0 1	
			NCI	40	0 1 2	
			"Pkt	41-44		
			NCII	45	0 1 2	
			"Pkt	46-49		
			9k	54		
			E.Nr.	55-56		
			Schiz. Syndrom			
			SS	57	0 1	
			RDC/SZ	58	0 1 2	
			RDC/SAP	59	0 1 2	
			DSM/SZ	60	0 1	
			DSM/SF	61	0 1	
			DSM/RP	62	0 1	
			DSM/PP	63	0 1	
			FDC/SZ	64	0 1	
			Man. Syndrom			
			MS	65	0 1	
			RDC 5	66	0 1 2	
			RDC 6	67	0 1 2	
			DSM 296x	68	0 1	
			DSM 301.3	69	0 1	
			DSM 267	70	0 1	
			FDC 1.	71	0 1	
			FDC 1.1	72	0 1	
			Sch aff. Syndrom			
			RDC 2	73	0 1 2	
			RDC 3	74	0 1 2	

			Karte Ende		
--	--	--	------------	--	--

			Karte Ende		
--	--	--	------------	--	--

			Karte Ende		
--	--	--	------------	--	--

			Karte Ende		
--	--	--	------------	--	--

			Karte Ende		
--	--	--	------------	--	--

			Karte Ende		
--	--	--	------------	--	--

			Karte Ende		
--	--	--	------------	--	--

			Karte Ende		
--	--	--	------------	--	--

			Karte Ende		
--	--	--	------------	--	--

			Karte Ende		
--	--	--	------------	--	--

			Karte Ende		
--	--	--	------------	--	--

			Karte Ende		
--	--	--	------------	--	--

			Karte Ende		
--	--	--	------------	--	--

			Karte Ende		
--	--	--	------------	--	--

			Karte Ende		
--	--	--	------------	--	--

			Karte Ende		
--	--	--	------------	--	--

			Karte Ende		
--	--	--	------------	--	--

			Karte Ende		
--	--	--	------------	--	--

			Karte Ende		
--	--	--	------------	--	--

			Karte Ende		
--	--	--	------------	--	--

			Karte Ende		
--	--	--	------------	--	--

			Karte Ende		
--	--	--	------------	--	--

			Karte Ende		
--	--	--	------------	--	--

			Karte Ende		
--	--	--	------------	--	--

			Karte Ende		
--	--	--	------------	--	--

			Karte Ende		
--	--	--	------------	--	--

			Karte Ende		
--	--	--	------------	--	--

			Karte Ende		
--	--	--	------------	--	--

			Karte Ende		
--	--	--	------------	--	--

			Karte Ende		
--	--	--	------------	--	--

			Karte Ende		
--	--	--	------------	--	--

			Karte Ende		
--	--	--	------------	--	--

			Karte Ende		
--	--	--	------------	--	--

			Karte Ende		
--	--	--	------------	--	--

			Karte Ende		
--	--	--	------------	--	--

			Karte Ende		
--	--	--	------------	--	--

			Karte Ende		
--	--	--	------------	--	--

			Karte Ende		
--	--	--	------------	--	--

			Karte Ende		
--	--	--	------------	--	--

			Karte Ende		
--	--	--	------------	--	--

			Karte Ende		
--	--	--	------------	--	--

			Karte Ende		
--	--	--	------------	--	--

			Karte Ende		
--	--	--	------------	--	--

			Karte Ende		
--	--	--	------------	--	--

			Karte Ende		
--	--	--	------------	--	--

			Karte Ende		
--	--	--	------------	--	--

			Karte Ende		
--	--	--	------------	--	--

			Karte Ende		
--	--	--	------------	--	--

			Karte Ende		
--	--	--	------------	--	--

			Karte Ende		
--	--	--	------------	--	--

			Karte Ende		
--	--	--	------------	--	--

			Karte Ende		
--	--	--	------------	--	--

			Karte Ende		
--	--	--	------------	--	--

			Karte Ende		
--	--	--	------------	--	--

			Karte Ende		
--	--	--	------------	--	--

			Karte Ende		
--	--	--	------------	--	--

			Karte Ende		
--	--	--	------------	--	--

			Karte Ende		
--	--	--	------------	--	--

			Karte Ende		
--	--	--	------------	--	--

			Karte Ende		
--	--	--	------------	--	--

			Karte Ende		
--	--	--	------------	--	--

			Karte Ende		
--	--	--	------------	--	--

			Karte Ende		
--	--	--	------------	--	--

			Karte Ende		
--	--	--	------------	--	--

			Karte Ende		
--	--	--	------------	--	--

			Karte Ende		
--	--	--	------------	--	--

			Karte Ende		
--	--	--	------------	--	--

			Karte Ende		
--	--	--	------------	--	--

			Karte Ende		
--	--	--	------------	--	--

			Karte Ende		
--	--	--	------------	--	--

			Karte Ende		
--	--	--	------------	--	--

			Karte Ende		
--	--	--	------------	--	--

			Karte Ende		
--	--	--	------------	--	--

			Karte Ende		
--	--	--	------------	--	--

			Karte Ende		
--	--	--	------------	--	--

			Karte Ende		
--	--	--	------------	--	--

			Karte Ende		
--	--	--	------------	--	--

			Karte Ende		
--	--	--	------------	--	--

			Karte Ende		
--	--	--	------------	--	--

			Karte Ende		
--	--	--	------------	--	--

			Karte Ende		
--	--	--	------------	--	--

			Karte Ende		
--	--	--	------------	--	--

			Karte Ende		
--	--	--	------------	--	--

			Karte Ende		
--	--	--	------------	--	--

			Karte Ende		
--	--	--	------------	--	--

			Karte Ende		
--	--	--	------------	--	--

			Karte Ende		
--	--	--	------------	--	--

			Karte Ende		
--	--	--	------------	--	--

			Karte Ende		
--	--	--	------------	--	--

			Karte Ende		
--	--	--	------------	--	--

			Karte Ende		
--	--	--	------------	--	--

			Karte Ende		
--	--	--	------------	--	--

			Karte Ende		
--	--	--	------------	--	--

			Karte Ende		
--	--	--	------------	--	--

			Karte Ende		
--	--	--	------------	--	--

			Karte Ende		
--	--	--	------------	--	--

			Karte Ende		
--	--	--	------------	--	--

			Karte Ende		
--	--	--	------------	--	--

			Karte Ende		
--	--	--	------------	--	--

			Karte Ende		
--	--	--	------------	--	--

			Karte Ende		
--	--	--	------------	--	--

			Karte Ende		
--	--	--	------------	--	--

			Karte Ende		
--	--	--	------------	--	--

			Karte Ende		
--	--	--	------------	--	--

			Karte Ende		
--	--	--	------------	--	--

			Karte Ende		
--	--	--	------------	--	--

			Karte Ende		
--	--	--	------------	--	--

7. Literaturverzeichnis

- Abou-Saleh, M.T., Coppen, A.: Classification of depressive illness. Clinicopsychological correlates.
J. Aff. Disord. 6(1984)53-66
- American Psychiatric Association, Diagnostic and Statistical Manual of Mental Disorders, 3rd edition (DSM-III).
American Psychiatric Association, Washington DC, 1980
- American Psychiatric Association, Diagnostic and Statistical Manual of Mental Disorders, revised edition (DSM-III-R)
American Psychiatric Association, Washington DC, 1985
- Andreasen, N.C., Winokur, G.: Newer experimental methods for classifying depression.
Arch. Gen. Psychiatry 36 (1979) 447-452
- Andreasen, N.C., Grove, W.M., Maurer, R.: Cluster analysis and the classification of depression
Br. J. Psychiatry 137(1980)256-265
- Andreasen, N.C., McDonald-Scott, P., Grove, W.M., Keller, M.B., Shapiro, R.W., Hirschfeld, R.M.: Assessment of reliability in multicenter collaborative research with a videotape approach
Am. J. Psychiatry 139 (1982) 876-882
- Andreasen, N.C., Scheftner, W., Reich, Th., Hirschfeld, R.M., Endicott, J., Keller, M.B.: The validation of the concept of endogenous depression.
A family study approach
Arch. Gen. Psychiatry 43 (1986) 246-252
- Bech, P., Clemmesen, L.: The diagnosis of depression
- 20 years later
Acta Psychiat. Scand. Suppl. 310(1983)9-30
- Bech, P., Gjerris, A., Andersen, J., Boejholm, S., Kramp, P., Bolwig, T.G.: The Melancholia Scales and the Newcastle Scales: Item-combinations and inter-observer reliability.
Brit. J. Psychiat. 143 (1983) 58-63
- Behar, D., Winokur, G., Van Valkenburg, C.: Familial subtypes of depression: A clinical view
J. Clin. Psychiatry 41 (1980) 52-56
- Berner, P., Katschnig, H., Lenz, G.: The polydiagnostic approach: A method to clarify incongruences between different classification systems of the functional psychoses
Psychiat. J. University of Ottawa 7(1982)244-248
- Berner, P., Katschnig, H.: Improving the quality of diagnostic systems: multiaxiality and the polydiagnostic approach.
In: Pichot, P., Berner, P., Wolf, R., Thau, K.T. (Hrsg.): Psychiatry. The state of the art. Vol. 1: Clinical Psychopathology, Nomenclature and Classification, S. 31-36, Plenum Press, New York

- Berner, P., Kieffer, W., Musalek, M., Pakesch, G., Simhandl, Ch., Wancata, J.: Approaches to the assessment of schizophrenia in Europe. *Pharmacopsychiat.* 19(1986)33-36
- Boyer, P., Pull, C.B., Dreyfuss, I.J., Pichot, P.: A computerized diagnostic system for comparing alternative classification schemes of depression. *J.Aff.Disord.* 7 (1984) 59-171
- Bronisch, T., Schmid, W., Cranach, M.v.: Die diagnostische Reliabilität und Validität des PSE/CATEGO Systems *Arch Psychiatr. Nervenkr* 231(1982)269-281
- Carroll, B.J., Feinberg, M., Greden, J.F., Haskett, R.F., James, N.M., Steiner, M., Tarika, J.: Diagnosis of endogenous depression: comparison of clinical, research and neuroendocrine criteria. *J.Aff.Disord.* 2(1980)177-194
- Cicchetti, D.V., Sparrow, S.S.: Developing criteria for establishing the inter-rater-reliability of specific items in a given inventory: applications to the assessment of adaptive behaviours. *Am.J.Ment.Defic.* 86 (1981) 127-137
- Copeland, J.R.M.: Reactive and endogenous depressive illness and five- year outcome. *J.Aff.Disord.* 6 (1984) 153-162
- Copolov, D.L., Rubin, R.T., Mander, A.J., Sashidharan, S.P., Whitehouse, A.M.: DSM-III Melancholia: Do the criteria accurately and reliably distinguish endogenous pattern depression? *J.Aff.Disord.* 10(1986)191-202
- Craig, Th.J., Goodman, A.B., Haugland, G.: Impact of DSM-III on clinical practice. *Am. J. Psychiatry* 139(1982)922-925
- Davidson, J., Turnbull, C., Strickland, R., Belyea, M.: Comparative diagnostic criteria for melancholia and endogenous depression *Arch. Gen. Psychiatry* 41(1984a)506-511
- Davidson, J., Lipper, S., Zung, W.W., Strickland, R., Krishnan, R., Mahorney, S.: Validation of four definitions of melancholia by the dexamethasone suppression test *Am. J. Psychiatry* 141(1984b)1220-1223
- Diagnoseschlüssel und Glossar psychiatrischer Krankheiten, 5. Aufl., korrigiert nach der 9. Revision der ICD. Springer Verlag, 1980
- Endicott, J.E., Spitzer, R.L.: A diagnostic interview: the Schedule for Affective Disorders and Schizophrenia *Arch. Gen. Psychiatry* 35(1978)837-844
- Everitt, B.S.: Moments of the statistic kappa and weighted kappa. *Br.J.Math.Stat.Psychol.* 21(1968)97-103
- Feighner, J.P., Robins, E., Guze, S.B.: Diagnostic criteria for use in psychiatric research *Arch.Gen.Psychiat.* 26(1972)57-63

- Feinberg, M., Carroll, B.J.: Separation of subtypes of depression using discriminant analysis. I. Separation of unipolar endogenous depression from non-endogenous depression.
Brit. J. Psychiatry 140 (1982) 384-391
- Feinberg, M., Carroll, B.J.: Separation of bipolar endogenous depression from non-endogenous (neurotic) depression.
J. Aff. Disord. 5 (1983) 129-139
- Fleiss, J.L.: Measuring nominal scale agreement between many raters.
Psychol. Bull. 76 (1971) 378-382.
- Friedmann, R.C., Clarkin, J.F., Corn, R. et al.: DSM-III and affective pathology in hospitalized adolescents
J. Nerv. Ment. Dis. 170 (1982) 511-521
- Fuhrer, R., Rouillon, F., Lellouch, J.: Diagnostic reliability among French psychiatrists using DSM-III criteria
Acta Psychiatr. Scand. 73 (1986) 12-16
- Helzer, J.E., Robins, L.N., Taibleson, M.: Reliability of psychiatric diagnosis: I. A methodological review
Arch. Gen. Psychiatry 34 (1977a) 129-133
- Helzer, J.E., Clayton, P.J., Pambakian, R., Woodruff, R.A.: Reliability of psychiatric diagnosis: II. The test/retest reliability of diagnostic classification
Arch. Gen. Psychiatry 34 (1977b) 136-141
- Helzer, J.E., Clayton, P.J., Pambakian, R., Woodruff, R.A.: Concurrent diagnostic validity of a structured psychiatric interview
Arch. Gen. Psychiatry 35 (1978) 849-853
- Hyer, St.E., Williams, J.B.W., Spitzer, R.L.: Reliability in the DSM-III Field Trials: interview vs. case summary.
Arch. Gen. Psychiatry 39 (1982) 1275-1279
- Jampala, V.C., Sierles, F.S., Taylor, M.A.: Consumer's views of DSM-III: Attitudes and practices of US psychiatrists and 1984 graduating psychiatric residents
Am. J. Psychiatry 142 (1986) 148-153
- Katschnig, H., Seelig, P.: The polydiagnostic approach in research of depression. In: Pichot, P., Berner, P., Wolf, R., Thau, K.T. (Hrsgb): Psychiatry. The state of the art, Vol. 1: Clinical Psychopathology, Nomenclature and Classification, S. 109-114,
Plenum Press, New York
- Kendell, R.E.: The classification of depressions: a review of contemporary confusion.
Brit. J. Psychiatry 129 (1976) 15-28
- Kendell, R.E.: The choice of diagnostic criteria for biological research
Arch. Gen. Psychiatry 39 (1982) 1334-1339
- Kendell, R.E.: Reflections on psychiatric classification - for the architects of DSM-IV and ICD-10.
Integr. Psychiat. 2 (1984) 43-47

- Kupfer, D.J., Rush, A.J.: Recommendations for depression publications. Information for authors.
J.Nerv.Ment.Dis. 171(1983)459-460
- Maier, W., Philipp, M.: A polydiagnostic scale for dimensional classification of endogenous depression.
Derivation and validation.
Acta Psychiat.Scand. 73(1986a)112-121
- Maier, W., Philipp, M.: Construct validity of the DSM-III and RDC classification of Melancholia (endogenous depression).
J.Psychiat.Res. 20(1986b)289-299
- Maier, W., Philipp, M., Buller, R.: Computability between ICD-9 and DSM-III classification of endogenous depression (Melancholia).
Pharmacopsychiatry 19(1986c)58-61
- Maier, W., Philipp, M., Buller, R., Benkert, O.: Sources of disagreement between clinical (ICD-9) and operational (DSM-III, RDC) diagnosis of endogenous depression (Melancholia).
J.Aff.Disord. 11(1986d)235-243
- Maier, W., Philipp, M., Heuser, I.: Dimensional assessment of endogenous depression based on a polydiagnostic approach.
Psychopathology 19(1986e)267-275
- Maier, W., Philipp, M., Schlegel, S., Heuser, I., Buller, R., Frommberger, U., Wetzel, H., Wilhelmi, D., Wittenborg, A., Demuth, W.: Operational diagnoses for schizophrenic and schizoaffective disorders.
I. Inter-rater-reliability.
Pharmacopsychiatry 19(1986f)178-179
- Maier, W., Philipp, M., Schlegel, S., Buller, R.: Reliability and validity of the Newcastle Scales in relation to ICD-9 classification.
Acta Psychiat.Scand. 76 (1987a) 619-627
- Maier, W., Heuser, I., Philipp, M., Schlegel, S., Benkert, O.: Diagnostic determinants of short-term spontaneous course of Major Depression (Letter to the editors)
Psychiatry Research 23 (1987b) 117-118
- Maier, W., Philipp, M., Heuser, I., Schlegel, S., Buller, R.: Improving depression severity assessment:
I. Reliability, internal validity and sensitivity to change of three observer depression scales
J.Psychiat.Res.22 (1988a) 3-12
- Maier, W., Heuser, I., Philipp, M., Frommberger, U., Demuth, W.: Improving depression severity assessment:
II. Content, concurrent and external validity of three observer depression scales
J.Psychiat.Res.22 (1988b) 13-19
- Maier, W., Buller, R., Philipp, M., Heuser, I.: The Hamilton anxiety scale: reliability, validity and sensitivity to change in anxiety and depressive disorders
J. Aff. Disorders 14 (1988c) 61-68

- Marneros, A., Deister, A., Rohde, A., Jünemann, H., Fimmers, R.:
Long-term course of schizoaffective disorders.
Part I: Definitions, methods, frequency of episodes
and cycles
Eur. Arch. Psychiatr. Neurol. Sci. 237 (1988) 264-275
- Marneros, A., Deister, A., Rohde, A., Jünemann, H., Fimmers, R.:
Long-term course of schizoaffective disorders.
Part II: Length of cycles, episodes and intervals
Eur. Arch. Psychiatr. Neurol. Sci. 237 (1988) 276-282
- Marneros, A., Deister, A., Rohde, A., Fimmers, R., Jünemann, H.:
Long-term course of schizoaffective disorders.
Part III: Onset, type of episodes and syndrome shift,
precipitating factors, suicidality, seasonality,
inactivity of illness, and outcome
Eur. Arch. Psychiatr. Neurol. Sci. 237 (1988) 283-290
- Maurer, M., Kuny, S., Woggon, B., Dittrich, A., von Luckner, N.:
Comparison of the AMP system and the CPRS with regard
to interrater reliability.
Neuropsychobiol. 12(1984)27-33
- Möller, H.-J., Hohe-Schramm, M., v. Zerssen, D.:
Operationalisierte Diagnostik endogener Psychosen nach
RDC und DSM-III. bereinstimmung mit der ICD-Diagnostik
und prognostische Bedeutung für der längerfristigen
Krankheitsverlauf.
in: Beckmann, H., Laux, G. (Hrsg.): Biologische
Psychiatrie - Synopsis 1986/87, Springer Verlag 1988
- Nelson, J.C., Charney, D.S.: Primary affective disorder and
the endogenous-reactive distinction
Arch. Gen. Psychiatry 37(1980)787-793
- Nelson, J.C., Charney, D.S.: The symptoms of Major Depressive
Illness
Am. J. Psychiatry 138 (1981a) 1-14
- Nelson, J.C., Charney, D.S., Quinlan, D.M.: Evaluation of the
DSM-III criteria for melancholia
Arch. Gen. Psychiatry 38(1981b)555-559
- Philipp, M.: Klassifikation manischer und depressiver
Psychosen.
In: U.H. Peters: Die Psychologie des XX. Jahrhunderts.
Bd. X: Ergebnisse für die Medizin (2), S. 471-478.
Kindler-Verlag, Zürich 1980
- Philipp, M., Maier, W.: Operational diagnosis of endogenous
depression: I. Comparison with clinical diagnosis.
Pharmacopsychiat. 18(1985a)112-113
- Philipp, M., Maier, W., Benkert, O.: Operational diagnosis of
endogenous depression II. Comparison of 8 different
operational diagnoses.
Psychopathol. 18(1985b)218-225
- Philipp, M., Maier, W.: The relevance of categorial and
dimensional classification systems for the
comparability of patient samples in psycho-
pharmacological research of depression.
Pharmacopsychiat. 19(1986a)43-47

- Philipp,M., Maier,W.: The polydiagnostic interview:
A structured interview for polydiagnostic
classification of psychiatric patients.
Psychopathol. 19(1986b)175-185
- Philipp,M., Maier,W.: Methodological problems in
polydiagnostic assessment of endogenous depression.
Psychopathol. 19(1986c)236-243
- Philipp,M., Maier,W., Frommberger,U., Wilhelmi,D.,
Wetzel,H., Schlegel,S., Buller,R., Heuser,I.,
Demuth,W., Wittenborg,A.: Operational diagnoses for
schizophrenic and schizoaffective disorders:
II. Association between diagnostic systems.
Pharmacopsychiat. 19(1986d)180-181
- Philipp,M., Maier,W.: Diagnosesysteme endogener
Depressionen
Springer Verlag, Berlin - Heidelberg 1987
- Pull,C.B., Pichot,P., Overall,J.E.: Comparison of French
and American diagnostic practices for schizophrenia
and affective disorders.
Psychopharmacol.Bull. 16(1980)43-45
- Robins,L.N., Helzer,J.E., Croughan,J., Ratcliff,K.S.:
National Institute of Mental Health Diagnostic
Interview Schedule: its history, characteristics
and validity.
Arch. Gen. Psychiatry 38(1981)381-389
- Sachs,L.: Angewandte Statistik. 4.Auflage.
Springer Verlag (1974)
- Silverstein,M.L., Warren,R.A., Harrow,M., Grinker,R.R.:
Changes in diagnosis from DSM-II to the Research
Diagnostic Criteria and DSM-III
Am. J. Psychiatry 139(1982)366-368
- Singerman,B., Stolzman,R.K., Robins,L.N., Helzer,J.E.,
Croughan, J.: Diagnostic concordance between
DSM-III, Feighner and RDC
J. Clin. Psychiatry 42(1981)422-426
- Spitzer,R.L., Fleiss J.L.: A re-analysis of the reliability
of psychiatric diagnosis.
Br.J.Psychiat. 125 (1974)341-347
- Spitzer,R.L., Endicott,J., Robins,E.: Research Diagnostic
Criteria Psychopharmacol.Bull. 11(1975)22-25
- Spitzer,R.L., Endicott,J., Robins,E.: Research Diagnostic
Criteria Rationale and Reliability
Arch. Gen. Psychiatry 35(1978)773-782
- Spitzer,R.L., Forman,J., Nee,J.: DSM-III Field Trials:
I. Initial Interrater Diagnostic Reliability
Am. J. Psychiatry 136(1979)815-817
- Spitzer,R.L., Williams, J.B.W.: The DSM-III classification
of affective disorders.
Acta Psychiatr. Scandinavica Suppl.310(1983a)106-116
- Spitzer,R.L.: Psychiatric diagnosis: Are clinicians still
necessary?
Compr. Psychiatry 24(1983b)399-411

- Stewart, J.W., Quitkin, F.M., Liebowitz, M.R., McGrath, P.J., Klein, D.M.: Efficacy of Desipramine in depressed outpatients. Response according to Research Diagnostic Criteria Diagnoses and severity of illness
Arch. Gen. Psychiatry 40(1983)202-207
- Uebersax, J.S.: A design-independent method for measuring the reliability of psychiatric diagnosis
J. Psychiatr. Res. 17(1982/83)335-342
- Williams, J.B.W., Spitzer, R.L.: Research Diagnostic Criteria and DSM-III: an annotated comparison
Arch. Gen. Psychiatry 39(1982)1283-1289
- Wing, J.K., Cooper, J.E., Sartorius, N.: Measurement and classification of psychiatric symptoms.
Cambridge University Press, Cambridge 1974
- Winokur, G.: The validity of Neurotic-Reactive Depression. New data and reappraisal
Arch. Gen. Psychiatry 42 (1985) 1116-1121
- Wittchen, H.U., Semler, G., v. Zerssen, D.: A comparison of two diagnostic methods: clinical ICD diagnoses vs DSM-III and RDC using the Diagnostic Interview Schedule
Arch. Gen. Psychiatry 42(1985)677-684
- Zimmerman, M., Stangl, D., Pfohl, B.M., Coryell, W.: Diagnosing RDC endogenous depression and DSM-III melancholia.
Psychiat. Res. 12(1984)173-174
- Zimmerman, M., Coryell, W.: Validity of DSM-III and RDC
Arch. Gen. Psychiatry 42(1985a)633-634
- Zimmerman, M., Coryell, W., Pfohl, B.M., Stangl, D.: Four definitions of endogenous depression and the dexamethasone suppression test.
J. Aff. Disord. 8(1985b)37-45
- Zimmerman, M., Coryell, W., Pfohl, B.M.: The treatment validity of DSM-III melancholic subtyping.
Psychiat. Res. 16(1985c)37-43
- Zimmerman, M., Coryell, W., Pfohl, B.M., Stangl, D.: The validity of four definitions of endogenous depression. II. Clinical, demographical familial and psychosocial correlates.
Arch. Gen. Psychiat. 43(1986a)234-244
- Zimmerman, M., Coryell, W., Pfohl, B.M.: Melancholic subtyping: a qualitative or quantitative distinction?
Am. J. Psychiat. 143(1986b)98-100
- Zimmerman, M., Coryell, W., Pfohl, B.M., Stangl, D.: The reliability of the family history method for psychiatric diagnoses
Arch. Gen. Psychiatry 45 (1988) 320-322

Dieter Gerau

3384 Liebenburg
Schloßstr. 4
Tel. (05346) 2722

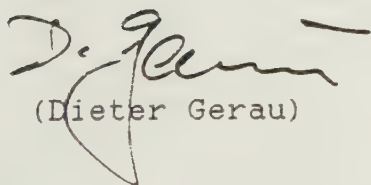
E R K L Ä R U N G

Ich versichere durch meine Unterschrift, daß die von mir dem Fachbereich Medizin der Johannes Gutenberg Universität Mainz vorgelegte Dissertation selbständig angefertigt worden ist.

Alle dazu benutzten Veröffentlichungen sind als solche im Text der Dissertation sowie im Literaturverzeichnis angegeben, wörtlich oder inhaltlich aus diesen Publikationen übernommene Textteile sind als solche kenntlich gemacht und nachgewiesen.

Mit der einzigen Ausnahme der Erstellung von Basic-Computer-Programmen habe ich keine weitere fremde Hilfe in Anspruch genommen.

Liebenburg, im Juli 1989


(Dieter Gerau)

Danksagung

Mein Dank gilt zunächst Herrn Prof. Dr. Michael Philipp, der mir nicht nur das Thema der vorliegenden Arbeit zur Verfügung stellte, sondern mich auch während der langen Entstehungsphase dieser Dissertation immer wieder durch Rat, Kritik und fachliche Hinweise unterstützt und motiviert hat, und dem ich meine ersten eingehenderen Kenntnisse des Faches Psychiatrie verdanke, von denen ich auch heute noch im täglichen Berufsleben zehre.

Weiterhin danke ich den Kollegen B.Diehl, U.Frommberger, I.Heuser und A.Sonntag, die im Jahre 1984 die Zweit-Ratings der von mir ausgewerteten Krankengeschichten angefertigt haben.

Dank schulde ich auch Herrn K.-P. Kossakowski, Hamburg, der mir bei der Erstellung der Basic-Programme, die ich zur statistischen Analyse benötigte, immer wieder behilflich war, und auf dessen Rat ich mich stets verlassen konnte.

Ich danke meiner Frau für ihre Geduld.

Lebenslauf

Name: Dieter Gerau
Geboren: 29.04.1958 in Hassloch/Pfalz
Konfession: evangelisch
Staatsangehörigkeit: deutsch
Wohnort: Schloßstraße 4, 3384 Liebenburg
Eltern: Erich-Wilhelm Gerau, Postbeamter
Helene Gerau, geb. Wink, Hausfrau
Familienstand: verheiratet,
Eheschließung am 23.09.1983
Ehefrau: Martina Gerau, geb. Mentrup
Krankenschwester, geb. 08.02.1961
Kinder: Philipp David Leonidas Gerau,
geb. 24.05.1989

Schulbildung: 1965-1969 Ernst-Reuter-Grundschule
Hassloch
1969-1977 Kurfürst-Ruprecht-Gymnasium
Neustadt/Weinstraße
Abitur: Juni 1977, Durchschnitt 1,6

Studium: ab April 1978 an der
Johannes-Gutenberg-Universität, Mainz
Fachbereich Humanmedizin
Ärztliche Vorprüfung: März 1980
1. Staatsexamen: März 1981
2. Staatsexamen: März 1983
Praktisches Jahr: Oktober 1983 bis September 1984
Universitätsklinikum Mainz
Innere Medizin, Chirurgie, Psychiatrie
3. Staatsexamen: Oktober 1984
Approbation: 22. November 1984

Berufstätigkeit: Januar 1985 bis Februar 1986
Stabsarzt, Bundeswehr
seit 1. März 1986
Assistenzarzt
Privatnervenklinik Dr. Fontheim
3384 Liebenburg

Liebenburg, den 27. Juli 1989

